OXFORD MATHEMATICAL MONOGRAPHS

General Relativity and the Einstein Equations

YVONNE CHOQUET-BRUHAT



OXFORD SCIENCE PUBLICATIONS

OXFORD MATHEMATICAL MONOGRAPHS

Series Editors

J. M. BALL W. T. GOWERS
N. J. HITCHIN L. NIRENBERG
R. PENROSE A. WILES

OXFORD MATHEMATICAL MONOGRAPHS

For a full of titles please visit
http://www.oup.co.uk/academic/science/maths/series/omm/
Donaldson and Kronheimer: *The geometry of four-mainifolds, paperback*
Woodhouse: *Geometric quantization, second edition, paperback*
Hirschfeld: *Projective geometries over finite fields, second edition*
Evans and Kawahigashi: *Quantum symmetries of operator algebras*
Klingen: *Arithmetical similarities: Prime decomposition and finite group theory*
Matsuzaki and Taniguchi: *Hyperbolic manifolds and Kleinian groups*
Macdonald: *Symmetric functions and Hall polynomials, second edition, paperback*
Catto, Le Bris, and Lions: *Mathematical theory of thermodynamic limits: Thomas-Fermi type models*
McDuff and Salamon: *Introduction to symplectic topology, paperback*
Holschneider: *Wavelets: An analysis tool, paperback*
Goldman: *Complex hyperbolic geometry*
Colbourn and Rosa: *Triple systems*
Kozlov, Maz'ya and Movchan: *Asymptotic analysis of fields in multi-structures*
Maugin: *Nonlinear waves in elastic crystals*
Dassios and Kleinman: *Low frequency scattering*
Ambrosio, Fusco and Pallara: *Functions of bounded variation and free discontinuity problems*
Slavyanov and Lay: *Special functions: A unified theory based on singularities*
Joyce: *Compact manifolds with special holonomy*
Carbone and Semmes: *A graphic apology for symmetry and implicitness*
Boos: *Classical and modern methods in summability*
Higson and Roe: *Analytic K-homology*
Semmes: *Some novel types of fractal geometry*
Iwaniec and Martin: *Geometric function theory and nonlinear analysis*
Johnson and Lapidus: *The Feynman integral and Feynman's operational calculus, paperback*
Lyons and Qian: *System control and rough paths*
Ranicki: *Algebraic and geometric surgery*
Ehrenpreis: *The radon transform*
Lennox and Robinson: *The theory of infinite soluble groups*
Ivanov: *The Fourth Janko Group*
Huybrechts: *Fourier-Mukai transforms in algebraic geometry*
Hida: *Hilbert modular forms and Iwasawa theory*
Boffi and Buchsbaum: *Threading homology through algebra*
Vazquez: *The Porous Medium Equation*
Benzoni-Gavage and Serre: *Multi-dimensional hyperbolic partial differential equations*
Calegari: *Foliations and the geometry of 3-manifolds*
Boyer and Galicki: *Sasakian Geometry*
Choquet-Bruhat: *General Relativity and the Einstein Equations*

General Relativity and Einstein's Equations

Yvonne Choquet-Bruhat

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Yvonne Choquet-Bruhat 2009

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2009

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn, Norfolk

ISBN 978–0–19–923072–3

1 3 5 7 9 10 8 6 4 2

Dedicated to the memory of my parents
Georges Bruhat, 1887–1944
who gave me the love of physics by his kindness and explanations, and after his death by his books.
Berthe Bruhat, born Hubert, 1892–1972
who always encouraged me to try to understand some of this world and to pursue scientific research.

Dedicated to my children; and grand children
Michelle, Daniel, Geneviève; and Raphael, Prisca, Abigail; Mahé, Mathurin, Enzo; Axel, Timothée, Andréas
who were and continue to be
my joy in life.

This page intentionally left blank

CONTENTS

I Lorentz geometry 1

1 Introduction 1
2 Manifolds 1
3 Differentiable mappings 2
4 Vectors and tensors 2
4.1 Tangent and cotangent space 2
4.2 Vector fields 3
4.3 Tensors and tensor fields 5
5 Pseudo-Riemannian metrics 6
5.1 General properties 6
5.2 Riemannian and Lorentzian metrics 8
6 Riemannian connection 9
7 Geodesics 12
8 Curvature 13
9 Geodesic deviation 14
10 Maximum of length and conjugate points 15
11 Linearized Ricci and Einstein tensors 17
12 Second derivative of the Ricci tensor 17

II Special Relativity 19

1 Newton's mechanics 19
1.1 The Galileo–Newton spacetime 19
1.2 Newton's dynamics – the Galileo group 20
2 Maxwell's equations 20
3 Minkowski spacetime 21
3.1 Definition 21
3.2 Maxwell's equations on M_4 22
4 Poincaré group 23
5 Lorentz group 24
5.1 General formulae 24
5.2 Transformation of electric and magnetic vector fields (case $n = 3$) 25
5.3 Lorentz contraction and dilatation 26
6 Special Relativity 26
6.1 Proper time 26
6.2 Proper frame and relative velocities 28
7 Dynamics of a pointlike mass 30
7.1 Newtonian law 30

7.2 Relativistic law 30
7.3 Equivalence of mass and energy 32
8 Continuous matter 33
8.1 Case of dust (incoherent matter) 34
8.2 Perfect fluids 35

III General relativity and Einstein's equations 37
1 Introduction 37
2 Newton's gravity law 37
3 General relativity 39
3.1 Physical motivations 39
4 Observations and experiments 40
4.1 Deviation of light rays 40
4.2 Proper time, gravitational time delay 40
5 Einstein's equations 42
5.1 Vacuum case 42
5.2 Equations with sources 43
6 Field sources 45
6.1 Electromagnetic sources 45
6.2 Electromagnetic potential 47
6.3 Yang–Mills fields 47
6.4 Scalar fields 49
6.5 Wave maps 49
6.6 Energy conditions 51
7 Lagrangians 51
7.1 Einstein–Hilbert Lagrangian 51
7.2 Lagrangians and stress energy tensors of sources 52
7.3 Coupled Lagrangian 53
8 Fluid sources 54
9 Einsteinian spacetimes 55
9.1 Definition 55
9.2 Regularity hypotheses 55
10 Newtonian approximation 57
10.1 Equations for potentials 57
10.2 Equations of motion 59
11 Gravitational waves 60
11.1 Minkowskian approximation 60
11.2 General linear waves 61
12 High-frequency gravitational waves 62
12.1 Phase and polarizations 64
12.2 Radiative coordinates 66
12.3 Energy conservation 68
13 Coupled electromagnetic and gravitational waves 68

13.1 Phase and polarizations 69
13.2 Propagation equations 69

IV Schwarzschild spacetime and black holes 72
1 Introduction 72
2 Spherically symmetric spacetimes 72
3 Schwarzschild metric 74
4 Other coordinates 75
4.1 Isotropic coordinates 75
4.2 Wave coordinates 76
4.3 Painlevé–Gullstrand-like coordinates 77
4.4 Regge–Wheeler coordinates 77
5 Schwarzschild spacetime 78
6 The motion of the planets and perihelion precession 78
6.1 Equations 78
6.2 Results of observations 81
6.3 Escape velocity 81
7 Stability of circular orbits 83
8 Deflection of light rays 84
8.1 Theoretical prediction 84
8.2 Results of observation 85
8.3 Fermat's principle and light travel parameter time 85
9 Red shift and time delay 86
10 Spherically symmetric interior solutions 87
10.1 Static solutions. Upper limit on mass 88
10.2 Matching with an exterior solution 91
10.3 Non-static solutions 91
11 The Schwarzschild black hole 92
11.1 The event horizon 92
11.2 The Eddington–Finkelstein extension 93
11.3 Eddington–Finkelstein white hole 94
11.4 Kruskal complete spacetime 94
11.5 Observations 96
12 Spherically symmetric gravitational collapse 96
12.1 Tolman metric 98
12.2 Monotonically decreasing density 101
13 The Reissner–Nordström solution 103
14 Schwarzschild spacetime in dimension $n + 1$ 104
14.1 Standard coordinates 104
14.2 Wave coordinates 104

V Cosmology 106
1 Introduction 106
2 Cosmological principle 107

3 Isotropic and homogeneous Riemannian manifolds 108
3.1 Isotropy 108
3.2 Homogeneity 109
4 Robertson–Walker spacetimes 111
4.1 Space metrics 112
4.2 Robertson–Walker spacetime metrics 113
4.3 Robertson–Walker dynamics 113
4.4 Einstein static universe 115
4.5 Cosmological red shift and the Hubble constant 115
4.6 De Sitter spacetime 118
4.7 Anti de Sitter (AdS) spacetime 120
5 Friedmann–Lemaître models. 121
5.1 Equation of state 121
5.2 General properties 122
5.3 Friedmann models 123
5.4 Some other models 124
5.5 Confrontation with observations 125
6 Homogeneous non-isotropic cosmologies 125
7 Bianchi class I universes 128
7.1 Kasner solutions 128
7.2 Models with matter 131
8 Bianchi type IX 132
9 The Kantowski–Sachs models 134
10 Taub and Taub NUT spacetimes 135
10.1 Taub spacetime 135
10.2 Taub NUT spacetime 136
11 Locally homogeneous models 136
11.1 n-dimensional compact manifolds 137
11.2 Compact 3-manifolds 139
12 Recent observations and conjectures 140

VI Local Cauchy problem 142
1 Introduction 142
2 Moving frame formulae 142
2.1 Frame and coframe 142
2.2 Metric 143
2.3 Connection 144
2.4 Curvature 145
3 $n+1$ splitting adapted to space slices 146
3.1 Adapted frame and coframe 146
3.2 Structure coefficients 147
3.3 Splitting of the connection. 147
3.4 Extrinsic curvature 148
3.5 Splitting of the Riemann tensor 148

4 Constraints and evolution 149
4.1 Equations. Conservation of constraints 149
5 Hamiltonian and symplectic formulation 151
5.1 Hamilton equations 151
5.2 Symplectic formulation 154
6 Cauchy problem 155
6.1 Definitions 155
6.2 The analytic case 156
7 Wave gauges 157
7.1 Wave coordinates 158
7.2 Generalized wave coordinates 161
7.3 Damped wave coordinates 162
7.4 Globalization in space, $\hat{e}$ wave gauges 162
7.5 Local in time existence in a wave gauge 164
8 Local existence for the full Einstein equations 166
8.1 Preservation of the wave gauges 166
8.2 Geometric local existence 168
8.3 Geometric uniqueness 168
8.4 Causality 170
9 Constraints in a wave gauge 172
10 Einstein equations with field sources 173
10.1 Maxwell constraints 174
10.2 Lorentz gauge 175
10.3 Existence and uniqueness theorems 176
10.4 Wave equation for F 177

VII Constraints 179
1 Introduction 179
2 Linearization and stability 180
2.1 Linearization of the constraints map, adjoint map 181
2.2 Linearization stability 183
3 CF (Conformally Formulated) constraints 186
3.1 Hamiltonian constraint 187
3.2 Momentum constraint 188
3.3 Scaling of the sources 189
3.4 Summary of results 194
3.5 Conformal transformation of the CF constraints 195
3.6 The momentum constraint as an elliptic system 197
4 Case $n = 2$ 200
5 Solutions on compact manifolds 200
6 Solution of the momentum constraint 201
7 Lichnerowicz equation 204
7.1 The Yamabe classification 204

7.2 Non-existence and uniqueness 210
7.3 Existence theorems 211
8 System of constraints 217
8.1 Constant mean curvature τ, sources with York-scaled momentum 217
8.2 Solutions with $\tau \not\equiv$ constant or $J_0 \not\equiv 0$ 218
9 Solutions on asymptotically Euclidean Manifolds 221
10 Momentum constraint 222
11 Solution of the Lichnerowicz equation 223
11.1 Uniqueness theorem 223
11.2 Generalized Brill–Cantor theorem 223
11.3 Existence theorems 226
12 Solutions of the system of constraints 229
12.1 Decoupled system 229
12.2 Coupled system 230
13 Gluing solutions of the constraint equations 232
13.1 Connected sum gluing 233
13.2 Exterior (Corvino–Schoen) gluing 235

VIII Other hyperbolic-elliptic well-posed systems 238
1 Introduction 238
2 Leray–Ohya non-hyperbolicity of ${}^{(4)}R_{ij} = 0$ 238
3 Wave equation for K 240
3.1 Hyperbolic system 240
3.2 Hyperbolic-elliptic system 242
4 Fourth-order non-strict and strict hyperbolic systems for $\bar{g}$ 243
5 First-order hyperbolic systems 243
5.1 FOSH systems 243
6 Bianchi–Einstein equations 244
6.1 Wave equation for the Riemann tensor 245
6.2 Case $n = 3$, FOS system 246
6.3 Cauchy-adapted frame 247
6.4 FOSH system for $u \equiv (E, H, D, B, \bar{g}, K, \bar{\Gamma})$ 250
6.5 Elliptic-hyperbolic system 250
7 Bel–Robinson tensor and energy 254
7.1 The Bel tensor 254
7.2 The Bel–Robinson tensor and energy 255
8 Bel–Robinson energy in a strip 256

IX Relativistic fluids 259
1 Introduction 259
2 Case of dust 260
2.1 Equations 260

2.2 Motion of isolated bodies (Choquet-Bruhat and Friedrichs 2006) 262
3 Charged dust 263
3.1 Equations 263
3.2 Existence and uniqueness theorem in wave and Lorentz gauges 264
3.3 Motion of isolated bodies 265
4 Perfect fluid, Euler equations 265
5 Energy properties 266
6 Particle number conservation 267
7 Thermodynamics 268
7.1 Definitions. Conservation of entropy 268
7.2 Equations of state 268
8 Wave fronts, propagation speeds, shocks 270
8.1 General definitions 270
8.2 Case of perfect fluids 272
8.3 Shocks 273
9 Stationary motion 274
10 Dynamic velocity for barotropic fluids 274
10.1 Fluid index and equations 274
10.2 Vorticity tensor and Helmholtz equations 276
10.3 Irrotational flows 277
11 General perfect fluids 278
12 Hyperbolic Leray system 279
12.1 Hyperbolicity of the Euler equations. 279
12.2 Reduced Einstein–Euler entropy system 280
12.3 Cauchy problem for the Einstein–Euler entropy system 281
12.4 Motion of isolated bodies 282
13 First-order symmetric hyperbolic system 282
14 Equations in a flow adapted frame 284
14.1 $n + 1$ splitting in a time adapted frame 285
14.2 Bianchi equations (case $n = 3$) 287
14.3 Vacuum case 287
14.4 Perfect fluid 288
14.5 Conclusion 290
15 Charged fluids 290
15.1 Equations 290
15.2 Fluids with zero conductivity 291
16 Fluids with finite conductivity 293
17 Magnetohydrodynamics 294
17.1 Equations 294
17.2 Wave fronts 295
18 Yang–Mills fluids 296

19 Dissipative fluids 297
19.1 Viscous fluids 297
19.2 The heat equation 300

X Relativistic kinetic theory 301
1 Introduction 301
2 Distribution function 302
2.1 Definition 302
2.2 Interpretation 303
2.3 Moments of the distribution function 304
3 Vlasov equations 307
3.1 Liouville–Vlasov equation. 307
3.2 Maxwell–Vlasov equation 310
3.3 Yang–Mills–Vlasov equation 311
3.4 Particles of a given rest mass 311
3.5 Conservation of moments 312
4 Cauchy problem for the Liouville–Vlasov equation 313
4.1 General solution 313
4.2 Distribution function on a Robertson–Walker space time 313
4.3 Energy estimates 314
4.4 Existence theorem 322
4.5 Stress energy tensor of a distribution function 323
5 The Einstein–Vlasov system 324
5.1 Constraints 324
5.2 Cauchy problem for the Einstein equations 325
5.3 Cauchy problem for the coupled system 325
6 The Einstein–Maxwell–Vlasov system 326
7 Boltzmann equation. Definitions 328
8 Moments and conservation laws 329
9 Einstein–Boltzmann system 331
10 Thermodynamics 331
10.1 Entropy and H theorem 331
10.2 Maxwell–Jüttner equilibrium distribution 333
10.3 Dissipative fluids 334
11 Extended thermodynamics 334
11.1 The phenomenological 14 moments theory 335
11.2 Extended thermodynamics of moments 338
11.3 Maximum characteristic velocity 339

XI Progressive waves 341
1 Introduction 341
2 Quasilinear systems 342
3 Quasilinear first-order systems 343
3.1 Phase and polarization 343
3.2 Propagation equations 344

4 Progressive waves in relativistic fluids 348
4.1 Equations 348
4.2 Progressive waves 348
4.3 Phases and polarizations 349
4.4 Polarization and propagation of acoustic waves 351
4.5 Polarization and propagation of matter waves 354
4.6 "Gauge" gravitational waves 354
5 Quasilinear quasidiagonal second-order systems 354
5.1 Definitions 354
5.2 Hyperquasilinear systems with f quadratic in Du 356
5.3 The null condition 358
6 Non quasidiagonal second-order systems 359
6.1 Phase and polarization 360
6.2 Propagation equations 360
7 Yang–Mills–scalar equations 361
7.1 Fields and equations 361
7.2 Phase and polarization 362
7.3 Propagation 363
8 Strong gravitational waves 364
8.1 Einstein equations 364
8.2 Phase and polarization 365
8.3 Propagation and back reaction 366
8.4 Example 368

XII Global hyperbolicity and causality 371
1 Introduction 371
2 Global existence of Lorentzian metrics 371
3 Time orientation 374
4 Futures and pasts 375
4.1 Paths and curves 375
4.2 Chronology and causality 375
5 Causal structure of Minkowski spacetime 377
6 Causal structures on general spacetimes 378
7 Geodesic coordinates, normal neighbourhoods 382
8 Topology on a space of paths 387
8.1 Rectifiable paths 387
8.2 Topology on sets of rectifiable paths 388
9 Global hyperbolicity 389
9.1 Definition and first criterion 389
9.2 Maximum of proper length 390
9.3 Images in V of subsets of $\mathcal{P}(x, y)$ 391
10 Strong and stable causalities 391
10.1 Strong causality 392
10.2 Stable causality 392

11 Cauchy surface 393
11.1 Domain of dependence. Cauchy horizon 393
11.2 Cauchy surface 394
11.3 Global existence of solutions of linear wave equations 397
11.4 Sufficient condition from analysis for global hyperbolicity 397
12 Globally hyperbolic Einsteinian spacetimes 399
12.1 Existence 399
12.2 Global uniqueness 399
12.3 Examples 400
13 Strong cosmic censorship 400

XIII Singularities 402
1 Introduction 402
2 Criteria for completeness or incompleteness 404
2.1 A completeness criterion 404
2.2 An incompleteness criterion 406
3 Congruence of timelike curves 408
3.1 Definitions 408
3.2 Geodesic deviation 410
3.3 Raychauduri equation 411
3.4 Null geodesic congruence 412
4 First singularity theorem 412
4.1 Conjugate points 412
4.2 Incompleteness theorem 414
5 Trapped surfaces and singularities 414
5.1 Trapped surfaces 414
5.2 Singularities linked to trapped surfaces 418
6 Black holes 418
6.1 Definitions 418
6.2 The Hawking area theorem 420
6.3 The Riemannian Penrose inequality. Case $n = 3$ 420
7 Weak cosmic censorship conjectures 421
7.1 Naked singularity 421
7.2 Weak cosmic censorship 422
8 Spherically symmetric Einstein scalar equations 423
8.1 Spherically symmetric spacetimes 423
8.2 Einstein equations in adapted frame 424
8.3 Reduction to one integro-differential equation 425
8.4 Bondi mass 427
8.5 Global existence for small data 427
8.6 Existence of a global generalized solution for large data 429

8.7 Structure of generalized solutions 432
8.8 Formation of a black hole. Cosmic censorship 433
8.9 Numerical results 434
8.10 Instability of naked singularities 434
9 Cosmological singularities. BKL conjecture 435
10 AVTD behaviour 441
10.1 Definitions 441
10.2 Fuchs theorem 441
11 Case of 1-parameter spatial isometry 443
11.1 Equations 443
11.2 VTD solutions of the $2+1$ Einstein evolution equations 445
11.3 The polarized case 446
11.4 The unpolarized case 449

XIV Stationary spacetimes and black holes 451
1 Introduction 451
2 Spacetimes with 1-parameter isometry group 452
2.1 Connection and Riemann tensor 454
2.2 Curvature tensor 454
2.3 Ricci tensor 455
3 Stationary spacetimes 455
3.1 General case 455
3.2 Static spacetimes 457
4 Gravitational solitons 458
4.1 Elementary proofs 458
4.2 Case $n = 3$, Komar mass 460
5 Electrovac solitons 465
6 The Einstein–Yang–Mills case 466
7 Stationary black holes 466
7.1 Definitions 466
7.2 Axisymmetry 467
8 The rigidity theorem for black holes 469
9 The Kerr metric and black hole 471
9.1 Kerr metric in Boyer–Linquist coordinates 471
9.2 The Kerr–Schild spacetime 472
10 Uniqueness of stationary black holes (dimension $3+1$) 474
10.1 Static black holes 475
10.2 Axisymmetric black holes 475
10.3 Uniqueness of the Kerr black hole 475
11 Further results 476
11.1 Multi black hole solutions 476
11.2 The Emparan–Reall "black rings" 478

XV Global existence theorems asymptotically Euclidean data 482
1 Introduction 482
2 Global existence for small data via the Penrose map 483
2.1 Yang–Mills and associated equations 484
2.2 Quasi-linear wave equations 484
2.3 Cases $n = 3$, the null condition 487
2.4 Wave maps 488
3 H. Friedrich conformal system, $n + 1 = 4$ 488
3.1 Equations 488
3.2 Friedrich hyperbolic system 490
4 Einstein's equations in higher dimensions 491
4.1 Conformal mapping 491
4.2 Transformed equations 492
4.3 Local Cauchy problem in R_x^{n+1}, $n \geq 5$ and odd 494
4.4 Global Cauchy problem 496
4.5 Conclusion 497
5 Christodoulou–Klainerman theorem 497
5.1 CK main theorem 497
5.2 Local existence 499
5.3 Global existence 500
6 The Klainerman–Nicolo theorem 504
7 The Linblad–Rodnianski theorem 505
7.1 The Einstein equations in wave coordinates 506
7.2 Initial data 507
7.3 Unknowns and norms 507
7.4 LR theorem 508

XVI Global existence theorems the cosmological case 510
1 Introduction 510
2 Gowdy cosmological models 511
3 S^1 invariant Einsteinian universes, equations 514
3.1 Introduction 514
3.2 Definition 514
3.3 Equations 515
3.4 Twist potential 515
3.5 Wave map system 516
3.6 Three-dimensional Einstein equations 517
3.7 Teichmüller parameters. 520
4 S^1 invariant Einstein universes, Cauchy problem 521
4.1 Cauchy data 521
4.2 Construction of A when F is known 522

4.3 Local in time existence theorem 522
4.4 Global existence theorem 522
4.5 Future complete existence 526
4.6 Einstein–Maxwell–Higgs system 526
4.7 Conclusion 527
5 Andersson–Moncrief theorem 528
5.1 CMC gauge, elliptic system for N and K 529
5.2 SH gauge, elliptic system for $\bar{g}$ and β 529
5.3 The Bianchi equations 530
5.4 Existence theorems 531
5.5 Global existence theorem 532
6 Einstein non-linear scalar field system 533

APPENDICES

I Sobolev spaces on Riemannian manifolds 534
1 Definitions 534
2 Embedding and multiplication properties 535
2.1 Open subsets of R^n 535
2.2 Riemannian manifolds 536
3 Weighted Sobolev spaces 537
3.1 Definitions 537
3.2 Embedding and multiplication properties 538

II Second-order elliptic systems on Riemannian manifolds 542
1 Linear elliptic systems 542
2 Linear elliptic systems on compact M 544
2.1 General second-order systems 544
2.2 Poisson operator 551
2.3 Conformal Laplace operator 552
3 Asymptotically Euclidean manifolds 553
3.1 Definitions 553
3.2 Second-order linear elliptic systems 554
4 Special systems 560
4.1 Poisson operator 560
4.2 Conformal Laplace operator 561
5 Equation $\Delta_\gamma \varphi = f(x, \varphi)$, compact M 562
6 $\Delta_\gamma \varphi = f(x, \varphi)$ *on* (M, γ) asymptotically Euclidean 567

III Quasi-diagonal, quasi-linear, second-order hyperbolic systems 571
1 Introduction 571
2 Wave equation on (V,g) 571
2.1 Definitions 572

2.2 Stress energy tensor. Energy momentum vector 572
2.3 Energy density 574
2.4 Energy equality on a compact domain 575
2.5 Energy inequality in a compact causal domain 578
2.6 Uniqueness theorem and causality 580
2.7 Case of a strip 581
2.8 Estimate of u 582
2.9 Cauchy problem 583
2.10 Generalizations 589
3 Quasidiagonal linear systems 590
3.1 Definitions 590
3.2 Stress energy tensor 590
3.3 Energy inequality in a compact causal domain 591
3.4 Case of a strip V_T 592
3.5 Existence, uniqueness, causality and continuity 594
3.6 Higher order estimates 594
3.7 Other hypotheses on g 600
3.8 Cauchy data in local spaces 601
4 Quasilinear systems 604
4.1 Semilinear systems 604
4.2 Further results for semilinear equations 607
4.3 Quasilinear systems 608
5 Global existence 613
5.1 Semilinear systems 613
5.2 Quasilinear equations 616

IV General hyperbolic systems 617
1 Introduction 617
2 Leray hyperbolic systems 617
2.1 Case of one equation 617
2.2 Leray hyperbolic systems 622
3 Leray–Ohya hyperbolic systems 624
4 First-order symmetric hyperbolic systems 625
4.1 FOSH systems on R^{n+1} 625
4.2 FOSH systems on a sliced manifold 627

V Cauchy–Kovalevski and Fuchs theorems 631
1 Introduction 631
2 Cauchy–Kovalevski theorem 631
2.1 Linear system 631
2.2 Non-linear system 634
3 Fuchs theorem 634
3.1 Definitions 634
3.2 Theorem 636

3.3 Equivalence with an integral equation 637
3.4 Equivalence with another mapping 638
3.5 Convergence of iterations 641
3.6 Global in space theorem 642

VI Conformal methods 643
1 Introduction 643
2 Conformal metrics. Confomorphisms 643
2.1 Connections of conformal metrics 643
2.2 Riemann tensors of conformal metrics 644
2.3 Ricci tensors of conformal metrics 644
3 The Weyl tensor 645
4 Conformal transformations of field equations 646
4.1 Maxwell and Yang–Mills equations 646
5 Invariance of wave equations 647
6 Penrose transform 647
7 Einstein spaces with cosmological constant 650
7.1 Conformal transformation of De Sitter spacetime 650
7.2 Conformal transformation of anti-De Sitter spacetime 650
8 Asymptotically simple spacetimes 650
8.1 Conformal compactifications 650
8.2 Black holes 652

VII Kaluza–Klein theories 653
1 Introduction 653
2 Isometries 653
3 Kaluza–Klein metrics 654
3.1 Metric in adapted frame 654
3.2 Structure coefficients 655
3.3 Kaluza–Klein connection 656
4 Curvature tensor 657
5 Ricci tensor and K–K equations 659
6 Equations in conformal spacetime metric 660

RELATED PAPERS

Causality of classical supergravity 665
Lecture Notes in Physics 1986, E. Flaherty ed. Springer, 61–84
Gravitation with gauss bonnet terms 689
Australian National University Publications, 1988 R. Bartnik ed. 53–72

Interaction of gravitational and fluid waves 709
In collaboration with A. Greco, Cericolo nat. di Paleruno 1994 Serie II n° 45, III. 123.
Positive-energy theorems 723
Relativity, Group and Topology II, B. Dewitt and R. Stora ed. 742–786.

REFERENCES 771

INDEX 781

FOREWORD

This book is essentially concerned with mathematical problems about the Einstein equations and their sources. However, I never forgot during the sixty years I worked on General Relativity (with some incursions into other subjects) that it is a physical theory. I remained interested in the interpretation of new observations and experiments which were made possible by the extraordinary progress in telescopes, satellites, particle colliders and nanotechnologies, and which confirmed more and more the physical relevance of General Relativity. I also remained fascinated by the remarkable theoretical developments which have considerably extended the scope of General Relativity as a theory of gravitation – as was in fact the hope of Einstein – and are a subject of active research. Lack of space and time, and also my incompetence, have prevented me from including these new experimental and theoretical fields of research in this book. The interested reader will find an overview and references in the recent article by T. Damour "General Relativity today"[1].

The first five chapters are an exposition of the classical foundations of General Relativity and Einstein's equations. They can form the text of a major undergraduate or first-year graduate course. They are mathematically precise, but they require only calculus-level knowledge in mathematics. They include the links with physics and astronomy which make General Relativity such an interesting subject. The following chapters are more advanced. Chapters VI–X prove global in space, local in time results for generic solutions of the Einstein equations and their classical sources. A reference for the mathematics used in these chapters are the books[2] *Analysis, Manifolds and Physics* written in collaboration with Cecile DeWitt. Useful results on partial differential equations, with proofs or sketch for them, are given in appendices. Chapter XI treats progressive waves, also called high-frequency waves, a precise and rigorous approximation method which is applied to the study of relativistic fluids and electromagnetic and gravitational waves. The last five chapters study global in time problems, where many conjectures are still open, although remarkable results have been obtained recently. Chapter XII gives fundamental definitions and proves basic results in global Lorentzian geometry relevant for General Relativity, while Chapter XIII

[1] Damour, T. (2007) General Relativity today, arXiv gr-qc 0704.0754, to appear in Proceedings Institute Poincaré seminar "Gravitation and experiments", Birkhaüser.

[2] Choquet-Bruhat, Y. and DeWitt Morette, C. *Analysis, Manifolds and Physics* I (1982) and II (1989), North Holland, referred to in the following as CB-DM1 and CB-DM2.

is concerned with singularities and Chapter XIV to the non-existence of gravitational solitons and the uniqueness of stationary black holes. The final chapters state, and sketch some proofs of, recently obtained global existence theorems for solutions of the Einstein equations.

ACKNOWLEDGEMENTS

The project of this book was made during research collaboration with James W. York. Unfortunately for me Jimmy was busy with other projects and my advancing years did not permit me to wait. To my regret I had to write this book alone; it would have been better with J. W. York as co-author. I thank him here for the pleasant and fruitful research collaboration we had in Chapel Hill and Cornell.

I am fortunate, however, that some eminent colleagues agreed to read parts of this book, and made very useful comments and corrections. I thank Piotr Chruściel, Spiros Cotsakis, Helmut Friedrich, Philippe LeFloch, Vincent Moncrief, Daniel Pollack, and Bernd Schmidt, who read some chapters or part of them, and particularly James Isenberg, who read carefully many chapters.

The book is enriched by some sections written by eminent specialists in the following chapters:

Gluing (Chapter VII) by James Isenberg and Robert Bartnik
Extended thermodynamics (Chapter X) by Tommaso Ruggeri
BKL conjecture (Chapter XIII) by Thibault Damour
Rigidity theorem (Chapter XIV) by James Isenberg
Further results on black holes (Chapter XIV) by Piotr Chruściel
Gowdy spacetimes (Chapter XV) by Vincent Moncrief

Antonio Greco permitted me to include the article we wrote in collaboration on the coupling of gravitational and fluid high-frequency waves.

I express to all these friends my warmest thanks.

My gratitude goes to the IHES, its Scientific Director, Jean-Pierre Bourguignon, and its permanent Professor, Thibault Damour, for providing me with a pleasant environment and working facilities, together with instructive and stimulating conversations without which I would not have been able to complete this book.

I thank the helpful secretaries of the IHES, and particularly Marie Claude Vergne who worked many hours to make the figures which adorn this book.

Notations

Equations III.7.2 denotes equation 7.2 in Chapter III
Section III.7.2 denotes Section 7.2 in Chapter III
when no roman number appears the reference concerns the current chapter
$:=$ is a definition, $\equiv$ is an identity.

This page intentionally left blank

I

LORENTZ GEOMETRY

1 Introduction

We give in this chapter a survey of the basic definitions of Riemannian and Lorentzian differential geometry that we will use in this book.

In the first nine sections we use the simplest formulations, in local coordinates, as they are needed for the first five chapters and physical applications. The later sections contain material used in the following, more advanced, chapters.

This chapter is a reminder to save the reader's time and to fix notations; it is not a course in Lorentzian geometry. More detailed and/or more abstract treatments can be found in many textbooks, in particular in Choquet-Bruhat and DeWitt-Morette[1].

2 Manifolds

The fundamental arena of differential geometry is a differentiable manifold. For the physicist the most concrete and useful definition makes apparent the local identification of a manifold with R^n, the space of sets of n real numbers with its usual topology. The definition proceeds as follows.

Recall that a mapping f between open sets of R^n, $f : u \to v$ is called a diffeomorphism if it is bijective and if f and its inverse mapping f^{-1} are differentiable. It is a C^k diffeomorphism if moreover f is of class C^k (f^{-1} is then also of class C^k).

A **chart** on a set X is a pair (U, ϕ), with U a subset of X, called a domain of the chart, and ϕ a bijection from U onto an open set u of R^n, i.e. an invertible mapping $\phi : U \to u$ by $x \mapsto \phi(x) \equiv (x^1, \ldots, x^n)$. The numbers x^i, $i = 1, \ldots, n$, are called **local coordinates** of the point $x \in X$.

An **atlas** on X is a collection of charts (U_I, ϕ_I), where $\{I\}$ is an arbitrary set of indices, whose domains cover X.

An atlas endows X with the structure of a **topological manifold**, of dimension n, if the mappings $\phi_I \circ \phi_J^{-1}$ are homeomorphisms (continuous bijections) between open sets of R^n, namely between $\phi_J(U_J \cap U_I)$ and $\phi_I(U_J \cap U_I)$. If these mappings are diffeomorphisms the manifold is a **differentiable manifold**, of class C^k if the mappings are of class C^k. The mappings $\phi_I \circ \phi_J^{-1}$ define changes

[1] Choquet-Bruhat, Y. and DeWitt-Morette, C. (1982) *Analysis Manifolds and Physics I and II* (revised edition 2000.), referred to subsequently as CB-DM1 and CB-DM2.

of local coordinates in the intersection $U_J \cap U_I$, $(x^1, \ldots, x^n) \mapsto (y^1, \ldots, y^n)$. The functions $y^i(x^1, \ldots, x^n)$ are of class C^k and the jacobian determinant $D(y)/D(x)$ with elements the partial derivatives $\partial x^i/\partial y^j$ is different from zero.

Two C^k atlases on X are said to be equivalent if their union is again a C^k atlas. One considers that they define the same C^k manifold.

The given definition of a manifold does not imply that it is a Hausdorf topological space, but admits the possibility for two points not to admit non-intersecting neighbourhoods. In the following we will, unless otherwise stated, consider only **Hausdorf manifolds** and call them simply manifolds.

An open set of R^n is oriented by the order of the coordinates $(x^1, \ldots, x^n)$. A differentiable manifold is said to be **orientable** (and oriented by the choice of coordinates) if its defining atlas is such that $D(y)/D(x) > 0$ in all intersections of domains of charts. Unless otherwise stated the manifolds considered will be connected and oriented.

3 Differentiable mappings

A function f on an n dimensional manifold V_n is a mapping $V_n \to R$ by $x \mapsto f(x)$. Its representative in local coordinates of the chart (U, ϕ) is a function on an open set of R^n, $f_\phi := f \circ \phi^{-1} : (x^i) \mapsto f(\phi^{-1}(x^i))$.

The function f is differentiable at x if f_ϕ is differentiable at $\phi^{-1}(x)$. This definition is chart independent if V_n is differentiable. The gradient, also called differential, of f is represented in a chart by the partial derivatives of f_ϕ. If (U, ϕ) and (U', ϕ') are two charts containing x, it holds that at x (calculus relations)[2]:

$$\frac{\partial f_\phi}{\partial x^i} = \frac{\partial f_{\phi'}}{\partial x^{j'}} \frac{\partial x^{j'}}{\partial x^i}. \tag{3.1}$$

This equivalence relation entitles the differential of f to the name of covariant vector (see below). A covariant vector is also called a 1-form. It is a geometric object, independent of a particular choice of coordinates.

A differentiable mapping f between differentiable manifolds, the source V_n and the target W_p, can be defined analogously. The differential at $x \in V_n$ is represented in a chart at $x \in V_n$ and a chart at $f(x) \in W_p$ by a linear mapping from R^n into R^p.

4 Vectors and tensors

4.1 *Tangent and cotangent space*

Tangent vectors and tensors on a differentiable manifold are geometric objects which can be intrinsically defined without the use of local coordinates. In particular, a tangent vector v_x at a point x of a manifold is the value at x of a linear first-order derivation operator acting on differentiable functions defined in

[2] By the Einstein convention, if the same index appears upstairs and downstairs in a term, it implies a summation over this index.

a neighbourhood of x (see for instance CB-DM1, III B 1). We recall here the definition through representatives.

A **tangent vector** v to a differentiable manifold V, at $x \in V$, is an equivalence class of triplets $(U_I, \phi_I, v_{\phi_I})$ where (U_I, ϕ_I) are charts containing x, while $v_{\phi_I} = (v^i_{\phi_I})$, $i = 1, \ldots, n$, are vectors in R^n. The equivalence relation is given by:

$$v^i_{\phi_I} = v^j_{\phi_J} \frac{\partial x^i_I}{\partial x^j_J} \tag{4.1}$$

where x^i_I and x^i_J are respectively local coordinates in the charts (U_I, ϕ_I) and (U_J, ϕ_J). The vector $v_\phi \in R^n$ is the **representative** of the vector v in the chart (U, ϕ). The vector v is attached to the manifold by the assumption, compatible with the equivalence relation, that the numbers v^i_ϕ are the components of v_ϕ in the frame of R^n defined by the tangent to the coordinate curves, where only one coordinate varies.

Tangent vectors at x constitute a vector space, the **tangent space** to V at x, denoted $T_x V$. An arbitrary set of n linearly independent tangent vectors constitute a frame at x. The **natural frame** associated to a chart (U, ϕ) is the set of n vectors $e_{(i)}$, $i = 1, \ldots, n$, such that $e^j_{(i),\phi} = \delta^j_i$. They are tangent vectors to the images in V of the coordinate curves of the chart. The numbers $v^i_{\phi_I}$ are the components[3] of the vector v in the natural frame.

The **cotangent space** T^*_x to V is the dual of T_x, i.e. the space of 1-forms on T_x. It is the vector space of **covariant vectors**. The components of a covariant vector at $x \in V_n$, in a chart (U, ϕ) containing x, is a set of n numbers ω_i, $i = 1, \ldots, n$. Under a change of chart from (U, ϕ) to (U, ϕ') it holds that

$$\omega_i = \omega_j \frac{\partial x^j}{\partial x'^i}. \tag{4.2}$$

Covariant vectors can also be defined by the equivalence relation (4.2). By (3.1), the differential at x of a differentiable function is a covariant vector.

A coframe is a set of n linearly independent covariant vectors. The **natural coframe** is the set of differentials dx^i of the coordinate functions $x \mapsto x^i$.

A frame and coframe, sets of n vectors $e_{(i)}$, $i, j = 1, \ldots, n$ and 1-forms $\theta^{(j)}$, $i, j = 1, \ldots, n$, are **dual frames** if

$$\theta^{(j)} e_{(i)} = \delta^j_i, \quad \text{the Kronecker symbol.} \tag{4.3}$$

4.2 *Vector fields*

A vector field [respectively a covariant vector field] on V assigns a tangent [respectively a covariant] vector at $x \in V$ to each point x.

[3] Traditionally components of contravariant vectors, usually called simply vectors, are written upstairs. Components of covariant vectors downstairs.

The relations (3.1) and (4.2) show that, given a differentiable function f on V, the quantity $v(f)$ defined for points x in the domain U of the chart (U, ϕ) by

$$v(f) := v^i_\phi \frac{\partial f_\phi}{\partial x^i_\phi} \tag{4.4}$$

is chart-independent: v defines a mapping between differentiable functions. It is easy to check that this mapping is linear:

$$v(f+g) = v(f) + v(g) \tag{4.5}$$

and satisfies the Leibniz rule:

$$v(fg) = fv(g) + gv(f) \tag{4.6}$$

The properties (4.6) and (4.5) characterize a **derivation** operator; $v(f)$ is called the derivation of f along v.

If we take for v a vector of a natural frame $e_{\phi,(i)}$, i.e. $v^j = \delta^j_i$ then

$$v(f) \equiv e_{\phi,(i)}(f) = \frac{\partial f(\phi^{-1})}{\partial x^i_\phi}. \tag{4.7}$$

More generally one calls **Pfaff derivative** and denotes by ∂_i the differential operator associated to the vector $e_{(i)}$ of an arbitrary frame. In the natural frame it coincides with the partial derivative.

The differential of a differentiable function f is a covariant vector field denoted df. It is called an exact 1 form.

One calls **frame** in an open set U of a manifold V_n a set of n vector fields which are linearly independent at each point of U. To each frame is associated by duality a coframe. The natural frame corresponds to the partial derivative operators $\partial/\partial x^i$, and the associated coframes are the 1-forms dx^i. Frames which are not the natural frame are usually called "moving frames" (see Chapter 6).

Manifolds do not always admit global frames, although three-dimensional orientable manifolds do.

Remark 4.1 *Vector fields, respectively 1-forms, on a manifold V_n admit a vector space structure, but are infinite dimensional. They are not linear combinations with constant coefficients of a finite numbers of them. For instance, such combinations of the 1-forms dx^i do not generate all 1-forms in the domain of a chart.*

Abstract index notation. In the course of this book we will often write formulae using representatives of vector fields (or tensor fields) in unspecified frames. These formulae are geometric ones bearing on geometric objects, equivalence classes of representatives. This procedure, used since the beginning of tensor calculus, has been baptized "abstract index notation" by Penrose to distinguish it from the case where coordinates with special properties are used. In the course of the text we mostly denote without indices the geometric objects, and we put indices in the formulas to make their reading easier.

4.3 *Tensors and tensor fields*

A covariant p-tensor at a point $x \in V$ can be defined as a p multilinear form on the p direct product of the tangent space T_xV. Contravariant tensors can be defined as multilinear forms on direct products of the cotangent space T_x^*V.

Tensors, contravariant or covariant can also be defined, like vectors, through equivalence relations between their components in various charts. For instance a covariant 2-tensor T at $x \in V$ is an equivalence class of triplets $(U_I, \phi_I, T_{\phi_I,ij})$, $i, j = 1, \ldots, n$, with the equivalence relation, by the law of change of components of T by change of local coordinates from (U, ϕ) to (U, ϕ') :

$$T_{i'j'} = \frac{\partial x^h}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{i'}} T_{hk}. \tag{4.8}$$

The space of covariant 2-tensors at x is denoted $T_x^* \otimes T_x^*$. The natural basis of this space associated to the chart with local coordinates x^i is denoted $dx^i \otimes dx^j$; that is T is given by:

$$T = T_{ij} dx^i \otimes dx^j. \tag{4.9}$$

where $dx^i \otimes dx^j$ is the covariant 2-tensor, bilinear form on $T_xV \times T_xV$, such that for any pair of vectors with components v^i, w^i in the natural frame it holds that:

$$(dx^i \otimes dx^j)(v, w) = v^i w^j \tag{4.10}$$

The **tensor product** $S \otimes T$ of a p tensor S and a q tensor T is intrinsically defined. It is the $p+q$ tensor with components defined by products of components. For example, the tensor product of a covariant vector ω with a contravariant 2-tensor T is the mixed 3-tensor $\omega \otimes T$ with components

$$(\omega \otimes T)_i{}^{jk} = \omega_i T^{jk}. \tag{4.11}$$

Although products of components are commutative, tensor products are *non-commutative*: $\omega \otimes T$ and $T \otimes \omega$ of the previous example do not belong to the same vector space.

The **contracted product** of a p contravariant tensor and a q covariant tensor is a tensor of order $p + q - 2$ whose components are obtained by summing over a repeated index appearing once upstairs and once downstairs.

The symmetry [respectively antisymmetry] property $T_{ij} = T_{ji}$, [respectively $T_{ij} = -T_{ji}$] are intrinsic properties (i.e independent of coordinates).

Tensor fields are assignments of a tensor at x to each point x of the manifold V. Differentiability can be defined in charts, and the notion of a C^k tensor field is chart-independent if the manifold is of class C^{k+1}.

The image of a contravariant tensor field under a differentiable mapping between differentiable manifolds $f: V \to W$ is not necessarily a tensor field on W unless f is a diffeomorphism.

The **pullback** f^* on V of a covariant tensor field on W is a covariant tensor field. For instance for a covariant vector ω and a mapping f given in local charts

by the functions $y^\alpha = f^\alpha(x^i)$, with y^α coordinates on W and x^i coordinates on V, it holds that

$$(f^*\omega)_i(x) = \frac{\partial y^\alpha}{\partial x^i}\omega_\alpha(y(x)). \tag{4.12}$$

The **Lie derivative** of a tensor field T with respect to a vector field X is a derivation operator from p-tensors into p-tensors. If f_t is the one parameter local group of diffeomorphism generated by X (see for instance CB-DM1, IIIC) the Lie derivative at $x \in V$ of a contravariant tensor is defined by

$$(\mathcal{L}_X T)(x) =: \lim_{t=0}[(f_t^{-1})'T(f_t(x)) - T(x)] \tag{4.13}$$

It is computed, for example for a contravariant 2-tensor, by

$$(\mathcal{L}_X T)^{jk} = X^i\frac{\partial T^{jk}}{\partial x^i} - T^{ik}\frac{\partial X^j}{\partial x^i} - T^{ji}\frac{\partial X^k}{\partial x^i}, \tag{4.14}$$

Analogous definitions give Lie derivatives of covariant tensors. For instance, for a covariant 2-tensor it holds that:

$$(\mathcal{L}_X T)_{jk} = X^i\frac{\partial T_{jk}}{\partial x^i} + T_{ik}\frac{\partial X^i}{\partial x^j} + T_{ji}\frac{\partial X^i}{\partial x^k}. \tag{4.15}$$

A totally antisymmetric covariant p-tensor field is also called an **exterior p-form**. A natural basis of the space of p forms is obtained by antisymmetrization. A p-form reads:

$$\omega = \frac{1}{p!}\omega_{i_1\ldots i_p}dx^{i_1}\wedge\cdots\wedge dx^{i_p}, \tag{4.16}$$

where

$$dx^i\wedge dx^j = -dx^j\wedge dx^i := \frac{1}{2}(dx^i\otimes dx^j - dx^j\otimes dx^i). \tag{4.17}$$

The **exterior derivative** of a p-form is a $p+1$-form given in local coordinates by

$$d\omega := \frac{1}{p!}d\omega_{i_1\ldots i_p}dx^{i_1}\wedge\cdots\wedge dx^{i_p}. \tag{4.18}$$

A form which is the differential of an exterior form is called an exact form. The exterior differential of an exact form is identically zero[4].

5 Pseudo-Riemannian metrics

5.1 *General properties*

A **pseudo-Riemannian metric** on a manifold V is a symmetric covariant 2-tensor field g such that the quadratic form it defines on contravariant vectors, $g(X,X)$, given in local charts by $g_{ij}X^iX^j$, is non-degenerate, that is the

[4] The reciprocal of this property is true only on simply connected manifolds (see e.g. CB-DM1 IV B 3).

determinant $Det(g)$ with elements g_{ij} does not vanish in any chart. This property is independent of the choice of the charts because under a change of local coordinates $(x'^i) \to (x^i)$ it holds that

$$Det(g) = Det(g') \left(\frac{D(x')}{D(x)} \right)^2. \tag{5.1}$$

The inverse of the matrix (g_{ij}) is denoted (g^{ij}) and defines the components of a contravariant symmetric 2-tensor. Through contracted products with the metric and its contravariant counterpart, one associates canonically contravariant and covariant tensors, for example

$$T_{ij} = g_{ih} g_{jk} T^{hk}. \tag{5.2}$$

A metric is traditionally written in the natural frame:

$$g = g_{ij} dx^i dx^j. \tag{5.3}$$

The **volume form** of the metric g is the exterior form which reads in local coordinates:

$$\mu_g = |\det g|^{\frac{1}{2}} dx^1 \wedge \cdots \wedge dx^n. \tag{5.4}$$

This exterior form induces, on domains oriented by the order $x^1, \ldots, x^n$, a **volume element**[5], often denoted by the same symbol,

$$\mu_g = |\det g|^{\frac{1}{2}} dx^1 \cdots dx^n. \tag{5.5}$$

If v is a tangent vector at x to V, the scalar

$$|g_x(v,v)|^{\frac{1}{2}} := |g_{ij}(x) v^i v^j|^{\frac{1}{2}} \tag{5.6}$$

is the **(pseudo) norm** (in g) of v. The **length** of a parametrized curve $\tau \mapsto x(\tau)$ joining two points of V with parameters τ_1 and τ_2 is:

$$l := \int_{\tau_1}^{\tau_2} \left| g_{ij}(x(\tau)) \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} \right|^{\frac{1}{2}} d\tau. \tag{5.7}$$

A curved $s \mapsto x(s)$ is parametrized by arc length s if

$$\left| g_{ij}(x(s)) \frac{dx^i}{ds} \frac{dx^j}{ds} \right| = 1. \tag{5.8}$$

An **isometry** of the pseudo-Riemannian manifold (V, g) is a diffeomorphism f which leaves g invariant; that is,

$$f^* g = g. \tag{5.9}$$

A metric is invariant by a 1-parameter group of isometries generated by the vector field X if its Lie derivative with respect to X vanishes.

[5] For integration of forms on manifolds see for instance CB-DM1, IV B 1.

Two pseudo-Riemannian manifolds (V, g) and (V', g') are called **locally isometric** if there exists a differentiable mapping f such that any point $x \in V$ admits a neighbourhood U, and $f(x)$ a neighbourhood U', with (U, g) and (U', g') isometric. Locally isometric manifolds have the same dimension but can have different topologies.

A pseudo-Riemannian manifold isometric with a pseudo-Euclidean space is called a **flat** space.

5.2 *Riemannian and Lorentzian metrics*

A metric is called **Riemannian** (or, for emphasis, properly Riemannian) if the quadratic form defined by g is positive definite.

A pseudo-Riemannian metric g is called a **Lorentzian metric** if the signature[6] of the quadratic form is $(-, +, +, \dots +)$. In the case of a Lorentzian manifold we denote its dimension by $n+1$ and we use Greek indices $\alpha, \beta, \dots, = 0, 1, 2, \dots, n$ for local coordinates or the representatives of vectors and tensors. In particular we write the metric in the natural frame:

$$g \equiv g_{\alpha\beta} dx^\alpha dx^\beta. \tag{5.10}$$

To say that g is Lorentzian is to say that this quadratic form admits a decomposition as a sum of squares of 1-forms θ^α with the following signs, independent[7] of the choice of these 1-forms,

$$g \equiv -(\theta^0)^2 + \sum_{i=1,\dots,n} (\theta^i)^2, \quad \theta^\alpha = a^\alpha{}_\beta dx^\beta. \tag{5.11}$$

The coframe (θ^α) and its dual are called **orthonormal Lorentzian frames**. Two such frames are deduced from each other by a Lorentz transformation (see Chapter 2).

The inequality in $T_x V$, the tangent space to V at x,

$$g_x(v, v) \leq 0, \quad v \in T_x v \tag{5.12}$$

defines a double cone C in $T_x V$, called the causal cone, with boundary the cone

$$g_x(v, v) = 0, \tag{5.13}$$

called the null cone, or the light cone.

The causal cone C splits into two convex cones, C_+ and C_-, characterized in an orthonormal Lorentzian frame by the properties $v \in C$ and:

$$C_+ : \quad \theta^0(v) \equiv v^0 > 0, \quad C_- : \quad v^0 < 0. \tag{5.14}$$

[6] We adopt the MTW convention. Some authors prefer the opposite convention, giving to Lorentzian metrics the signature $(+, -, -, \dots, -)$. Each of these conventions has its advantages and disadvantages, but of course they give equivalent geometrical results.

[7] By the properties of quadratic forms on vector spaces.

If it is possible to choose this splitting continuously on V the Lorentzian manifold is said to be **time orientable**. It is time oriented by the choice.

A vector $v \in T_x v$ is called **timelike** if it is such that

$$g_x(v, v) < 0 \tag{5.15}$$

It is called a **null** vector if

$$g_x(v, v) = 0 \tag{5.16}$$

and **causal** if

$$g_x(v, v) \leq 0. \tag{5.17}$$

A causal vector on a time-oriented Lorentzian manifold is **future** [respectively **past**] oriented if it is such that $v \in C_+$ [respectively $v \in C_-$].

A differentiable submanifold of codimension 1 is called spacelike [respectively, timelike, null] if its normal (in the metric g) is timelike [respectively, spacelike, null].

Remark 5.1 *A Lorentzian metric can always be written in a small enough neighbourhood by a change of coordinates under the form*

$$-N^2 dt^2 + g_{ij} dx^i dx^j. \tag{5.18}$$

Indeed, under a change of coordinates $(x'^\alpha) \mapsto (x^\beta)$ *with* $x^0 = x'^0$ *we have*

$$g'_{i0} = \frac{\partial x^j}{\partial x'^i} \left(g_{j0} + g_{jh} \frac{\partial x^h}{\partial x'^0} \right), \tag{5.19}$$

we make $g'_{i0} = 0$ *by solving the linear first-order system* $g_{j0} + g_{jh} \frac{\partial x^h}{\partial x'^0} = 0$ *for the functions* $x^h(x'^i, x'^0)$.

Global properties of Lorentzian manifolds are studied in Chapter 12.

6 Riemannian connection

Linear connection

Partial derivatives acting on components of tensor fields are not intrinsic operators on a manifold. Intrinsic derivation operators are defined by endowing the manifold with a new structure called a connection. These operators are called **covariant derivatives** and usually denoted ∇. They map differentiable vector fields into tensor fields and obey the following laws:

$$\nabla(v + w) = \nabla v + \nabla w \quad \text{linearity} \tag{6.1}$$

$$\nabla(fv) = f \nabla v + df \otimes v \quad \text{Leibnitz rule.} \tag{6.2}$$

Suppose that in an arbitrary frame, in the domain U of a chart, $v = v^i e_i$; then by the previous rules

$$\nabla v = v^i \nabla e_i + dv^i \otimes e_i \tag{6.3}$$

Hence ∇v is determined in U by the covariant derivatives of the basis vectors e_i, 2-tensor fields which we write in the chosen frame:

$$\nabla e_i = \omega_{ji}^h \theta^j \otimes e_h \tag{6.4}$$

The ω_{ij}^h are functions of the local coordinates on U. They are called **connection coefficients**.

The connection coefficients are not components of a tensor. Their transformation law under a change of local frame results from the tensorial character of ∇v and is found to be

$$\omega_{ji}^h = A_{h'}^h \partial_j A_i^{h'} + A_h^{h'} A_{j'}^j A_{i'}^i \omega_{j'i'}^{h'} \tag{6.5}$$

where A is the change of frame matrix, with elements $A_{i'}^i = \frac{\partial x^i}{\partial x^{i'}}$ in the case of change of natural frame.

The **covariant derivative** of a vector reads, by (6.3) and (6.4):

$$\nabla v = (\nabla_j v^i)\theta^j \otimes e_h, \quad \text{with} \quad \nabla_j v^i = \partial_j v^i + \omega_{ji}^h v^h. \tag{6.6}$$

The covariant derivative of v in the direction of another vector u is a vector $\nabla_u v$ with components

$$u^i \nabla_i v^j.$$

Remark 6.1 *The set of numbers $\nabla_i v^j$, for some fixed i, can be interpreted as the components of the vector which is the derivative of v in the direction of the frame vector $e_{(i)}$.*

The covariant derivative of a contravariant tensor in the direction of a vector u is a tensor of the same type as the derived one, and is defined in such a way that the derivative of a tensor product obeys the additivity and Leibnitz rules. The above remark makes the rule easy to apply on representatives[8].

The covariant derivative of a covariant vector is defined so that the covariant derivative of the scalar $v^i u_i$ is the ordinary derivative; therefore

$$\nabla_j u_i = \partial_j u_i - \omega_{ji}^h u_h. \tag{6.7}$$

These general definitions give for example

$$\nabla_i T_j{}^{hk}{}_l = \partial_i T_j{}^{hk}{}_l - \omega_{ij}^m T_m{}^{hk}{}_l + \omega_{im}^h T_j{}^{mk}{}_l + \omega_{im}^k T_j{}^{hm}{}_l - \omega_{il}^m T_j{}^{hk}{}_m. \tag{6.8}$$

Given a pseudo-Riemannian metric g, its **Riemannian connection** is a linear connection enjoying the following properties:

- It has no torsion i.e.:

$$\nabla(df) \quad \text{is a symmetric 2 tensor.} \tag{6.9}$$

[8] One cannot apply directly the Leibnitz rule to tensor fields because of the non-commutativity of the tensor product: $\nabla u \otimes v$ and $u \otimes \nabla v$ do not belong to the same vector space.

- The metric has covariant derivative zero:

$$\nabla g = 0. \tag{6.10}$$

The connection coefficients of a Riemannian connection in a natural frame are called **Christoffel symbols** and[9] denoted $\Gamma^{\lambda}_{\alpha\beta}$. Commutation of partial derivatives shows that (6.9) implies

$$\Gamma^{\lambda}_{\alpha\beta} = \Gamma^{\lambda}_{\beta\alpha}. \tag{6.11}$$

Equation (6.10) reads, with $\partial_\lambda := \frac{\partial}{\partial x^\lambda}$

$$\partial_\lambda g_{\alpha\beta} - \Gamma^{\mu}_{\lambda\alpha} g_{\mu\beta} - \Gamma^{\mu}_{\lambda\beta} g_{\alpha\mu} = 0. \tag{6.12}$$

An algebraic manipulation of (6.11) and (6.12) gives with $(g^{\alpha\beta})$ the inverse matrix of $g^{\alpha\beta}$,

$$\Gamma^{\lambda}_{\alpha\beta} = \frac{1}{2} g^{\lambda\mu}(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\alpha\mu} - \partial_\mu g_{\alpha\beta}). \tag{6.13}$$

One can show by using the transformation law (6.5) that, by a change of local coordinates, *the Christoffel symbols can be made to vanish at any one given point.*[10]

Exercise.

1. *Show that a Lie derivative can be expressed by replacing partial derivatives by covariant derivatives in an arbitrary Riemannian connection.*
2. *Show that the Lie derivative of a metric g with respect to a vector field can be written:*

$$(\mathcal{L}_g X)_{\alpha\beta} \equiv \nabla_\alpha X_\beta + \nabla_\alpha X_\beta. \tag{6.14}$$

Hint. Use the expressions (6.13) of the Christoffel symbols.

The **Hodge adjoint** $*\omega$ of a p-form ω on an $n+1$-dimensional spacetime is the $n+1-p$-form given by, with μ the volume form:

$$(*\omega)_{\alpha_{p+1}\ldots\alpha_n} := \frac{1}{p!}\mu_{\alpha_0\alpha_1\ldots\alpha_n}\omega^{\alpha_0\alpha_1\ldots\alpha_p} \tag{6.15}$$

The **codifferential operator** acting on exterior forms is the formal L^2 (in the metric g) adjoint of the operator d. It is defined, up to sign[11], by $\delta \equiv *d(*F)$. For a 2-form F it holds that:

$$(\delta F)^\beta \equiv \nabla_\alpha F^{\alpha\beta}. \tag{6.16}$$

[9] We start using Greek indices, in view of the application to metrics of Lorentzian signature.
[10] They can even be made to vanish along a given line.
[11] See for instance CB-DM1, V B 4.

7 Geodesics

Parallel transport. A vector field v is said to be parallelly transported along a curve $\lambda \mapsto x(\lambda)$ with tangent vector u if it satisfies along this curve the differential equation

$$u^\alpha \nabla_\alpha v^\beta = 0, \quad u^\alpha = \frac{dx^\alpha}{d\lambda} \tag{7.1}$$

A differentiable curve is called a **geodesic** if its tangent vector is parallelly transported

$$u^\alpha \nabla_\alpha u^\beta = 0; \tag{7.2}$$

that is,

$$\frac{dx^\alpha}{d\lambda}\left\{\frac{\partial u^\beta}{\partial x^\alpha} + \Gamma^\beta_{\alpha\lambda} u^\beta\right\} = 0 \tag{7.3}$$

equivalently

$$\frac{d^2x^\beta}{d\lambda^2} + \Gamma^\beta_{\alpha\gamma}\frac{dx^\alpha}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0. \tag{7.4}$$

The geodesic equation implies, by contraction with u_β, that $u^\alpha u_\alpha$ is constant along the curve, the timelike, null or space like character of a geodesic is the same all along. Moreover (see 5.7), the parameter λ is proportional to arc length if the curve is not a null curve. If the curve is a null curve the parameter λ is called the canonical parameter.

Theorem 7.1 *A time like geodesic joining two points of a pseudo-Riemannian manifold is a critical point of the length functional ℓ given by (5.7).*

Proof Consider for $C_\sigma\colon (\lambda, \sigma) \to x(\lambda, \sigma)$ a family of timelike curves near a timelike geodesic joining two points, parametrized proportionally to length

$$-g_{\alpha\beta}(x)\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \ell_\sigma^2 \text{ on } c_\sigma \tag{7.5}$$

The variation $\delta\ell_\sigma$ of the length ℓ_σ of c_σ is

$$\delta\ell_\sigma \equiv -\frac{\ell_\sigma^{-1}}{2}\int_0^1 \left\{\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\delta x^\lambda \frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} + 2g_{\alpha\beta}\frac{d\delta x^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right\} d\lambda \tag{7.6}$$

After integration by parts (with the $\delta x's$ vanishing at both ends $\lambda = 0$ and $\lambda = 1$), the expression of $\delta\ell$ becomes the following integral of a linear form in δx

$$\delta\ell \equiv \frac{\ell_\sigma^{-1}}{2}\int_0^1 \delta x^\alpha \left\{\frac{d}{d\lambda}\left(2g_{\alpha\beta}\frac{dx^\beta}{d\lambda} - \frac{\partial g_{\lambda\beta}}{\partial x^\alpha}\frac{dx^\lambda}{d\lambda}\frac{dx^\beta}{d\lambda}\right)\right\} d\lambda \tag{7.7}$$

i.e. with $u^\alpha = \frac{dx^\alpha}{d\lambda}$, hence $\frac{d}{d\lambda} = u^\alpha \frac{\partial}{\partial x^\alpha}$, and, using the expression of the Christoffel symbols,

$$\delta\ell_{\sigma_0} \equiv \ell_{\sigma_0}^{-1} \int_0^1 \delta x^\alpha \{u^\lambda \partial_\lambda u_\alpha - \Gamma^\beta_{\lambda\alpha} u^\lambda u_\beta\} d\lambda = 0 \tag{7.8}$$

which gives that $\delta\ell_{\sigma_0} = 0$ for all δx if and only if

$$u^\lambda \nabla_\lambda u_\alpha = 0. \tag{7.9}$$

□

8 Curvature

Curvature signals the non-commutativity of covariant derivatives. In non-infinitesimal terms it signals the non-identity of a vector and the vector parallelly transported along a closed loop.

The **curvature tensor** is defined as follows. Consider the commutation $(\nabla_\alpha \nabla_\beta - \nabla_\alpha \nabla_\beta) v^\lambda$ of two covariant derivatives of a vector v. It is a 2-covariant, 1-contravariant tensor which is found by calculus to depend linearly on v, and not on its derivatives: there exist coefficients $R_{\alpha\beta}{}^\lambda{}_\mu$ such that

$$(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) v^\lambda \equiv R_{\alpha\beta}{}^\lambda{}_\mu v^\mu \tag{8.1}$$

These coefficients are the components of a tensor[12], $Riem(g)$, antisymmetric in its first two indices, called the curvature tensor. The identity (8.1) is called the **Ricci identity**.

Using the formula

$$\nabla_\alpha v^\beta = \frac{\partial}{\partial x^\alpha} v^\beta + \Gamma^\beta_{\alpha\lambda} v^\lambda, \tag{8.2}$$

for the components in natural coordinates of the covariant derivative of a vector, the components of the Riemann tensor in such coordinates are found to be

$$R_{\alpha\beta}{}^\lambda{}_\mu = \partial_\alpha \Gamma^\lambda_{\beta\mu} - \partial_\beta \Gamma^\lambda_{\alpha\mu} + \Gamma^\lambda_{\alpha\rho} \Gamma^\rho_{\beta\mu} - \Gamma^\lambda_{\beta\rho} \Gamma^\rho_{\alpha\mu}. \tag{8.3}$$

It can be proved that the *vanishing of the Riemann tensor on a domain of a manifold implies that the metric is locally flat in this domain.*

The tensor $Riem(g)$ satisfies the **algebraic identities**, easily proved by taking coordinates such that Γ vanishes at the considered point

$$R_{\alpha\beta}{}^\lambda{}_\mu + R_{\mu\alpha}{}^\lambda{}_\beta + R_{\beta\mu}{}^\lambda{}_\alpha \equiv 0 \tag{8.4}$$

$$R_{\alpha\beta,\lambda\mu} \equiv R_{\lambda\mu,\alpha\beta} \tag{8.5}$$

Bianchi identities. The tensor $Riem(g)$ satisfies the following differential identities:

$$\nabla_\gamma R_{\alpha\beta}{}^\lambda{}_\mu + \nabla_\beta R_{\gamma\alpha}{}^\lambda{}_\mu + \nabla_\alpha R_{\beta\gamma}{}^\lambda{}_\mu \equiv 0. \tag{8.6}$$

[12] Because the left-hand side is a tensor and v is a vector.

Ricci tensor. The Ricci tensor, $Ricc(g)$, is the g trace of the curvature tensor, namely in components

$$R_{\alpha\beta} := R_{\lambda\alpha}{}^{\lambda}{}_{\beta} \equiv \partial_\lambda \Gamma^\lambda_{\alpha\beta} - \partial_\alpha \Gamma^\lambda_{\beta\lambda} + \Gamma^\lambda_{\alpha\beta}\Gamma^\mu_{\lambda\mu} - \Gamma^\lambda_{\alpha\mu}\Gamma^\mu_{\beta\lambda}. \tag{8.7}$$

The identity (8.5) shows that it is symmetric.

Scalar curvature. The scalar curvature is the g trace of the Ricci tensor:

$$R := g^{\alpha\beta} R_{\alpha\beta}. \tag{8.8}$$

Contracted Bianchi identities. The following identities play an important role in general relativity. Contracting (8.6) once gives

$$\nabla_\gamma R_{\beta\mu} - \nabla_\beta R_{\gamma\mu} + \nabla_\alpha R_{\beta\gamma}{}^{\alpha}{}_{\mu} \equiv 0 \tag{8.9}$$

Another contraction gives

$$\nabla_\gamma S^\gamma_\mu \equiv 0, \quad \text{with} \quad S_{\alpha\beta} =: R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R. \tag{8.10}$$

The tensor

$$S(g) := Ricc(g) - \frac{1}{2} g R \tag{8.11}$$

is called the **Einstein tensor**.

We will return to connections and curvature, this time in an arbitrary frame, in Chapter 6.

9 Geodesic deviation

Let C_σ be a 1-parameter family of timelike geodesics parametrized by their arc length s, $C_\sigma(s) := \psi(\sigma, s)$. Denote by $v = \partial\psi/\partial s$ the tangent vector to C_σ and by $X = \partial\psi/\partial\sigma$ the vector which characterizes the infinitesimal displacement of C_σ. These two vectors commute, i.e.

$$v^\alpha \nabla_\alpha X^\beta - X^\alpha \nabla_\alpha v^\beta = 0 \tag{9.1}$$

Derivating this relation in the direction of v gives, using the parallel transport of v,

$$v^\lambda v^\alpha \nabla_\lambda \nabla_\alpha X^\beta - v^\lambda \nabla_\lambda X^\alpha \nabla_\alpha v^\beta - X^\alpha v^\lambda \nabla_\lambda \nabla_\alpha v^\beta = 0 \tag{9.2}$$

Hence, using the Ricci identity

$$v^\lambda v^\alpha \nabla_\lambda \nabla_\alpha X^\beta - v^\lambda \nabla_\lambda X^\alpha \nabla_\alpha v^\beta - X^\alpha v^\lambda (\nabla_\alpha \nabla_\lambda v^\beta + R_{\lambda\alpha}{}^{\beta}{}_{\mu} v^\mu) = 0. \tag{9.3}$$

Hence, by the parallel transport of v

$$v^\lambda v^\alpha \nabla_\lambda \nabla_\alpha X^\beta - v^\lambda \nabla_\lambda X^\alpha \nabla_\alpha v^\beta + X^\alpha \nabla_\alpha v^\lambda \nabla_\lambda v^\beta - X^\alpha v^\lambda R_{\lambda\alpha}{}^{\beta}{}_{\mu} v^\mu] = 0. \tag{9.4}$$

The commutation (9.1) of v and X shows that this equation simplifies to:

$$v^\lambda v^\alpha \nabla_\lambda \nabla_\alpha X^\beta = X^\alpha v^\lambda v^\mu R_{\lambda\alpha}{}^{\beta}{}_{\mu} \tag{9.5}$$

also written

$$\nabla^2_{v^2} X^\beta \equiv \frac{D^2}{Ds^2} X^\beta = X^\alpha v^\lambda v^\mu R_{\lambda\alpha}{}^\beta{}_\mu \tag{9.6}$$

This equation, linking the curvature with the rate of acceleration of distance between nearby geodesics, is called the equation of **geodesic deviation**.

Remark 9.1 *Equation (9.6) also holds when v is the tangent vector to null geodesics canonically parametrized.*

10 Maximum of length and conjugate points

Consider the family of timelike curves $C_\sigma : \lambda \mapsto x(\lambda, \sigma)$ joining two points of a Lorentzian manifold. The second infinitesimal variation of the length ℓ at C_0 corresponding to the variation of α is by definition

$$\left\{ \frac{d^2 \ell_\sigma}{d\sigma^2} \right\}_{\sigma=0} \tag{10.1}$$

At a curve C_σ we have found

$$\frac{d\ell_\sigma}{d\sigma} \equiv \ell_\sigma^{-1} \int_0^1 (u_\sigma^\lambda \nabla_\lambda u_\sigma^\alpha) g_{\alpha\beta}(x) h_\sigma^\beta \, d\lambda \tag{10.2}$$

where u_σ is the vector $\partial x(\lambda, \sigma)/\partial\lambda$. We must consider the variation of the first derivative $d\ell/d\sigma$. We have the following formula for this second derivative

$$\frac{d^2 \ell\sigma}{d\sigma^2} = \frac{d}{d\sigma} \left\{ \ell_\sigma^{-1} \int_0^1 u_\sigma^\alpha \, \nabla_\alpha \, u_{\beta,\sigma} \, h_\sigma^\beta \, d\lambda \right\} \tag{10.3}$$

At a geodesic C_0 critical point of ℓ it holds that

$$u_0^\lambda \, \nabla_\lambda \, u_{\beta,0} = 0 \tag{10.4}$$

We deduce from the expression of $d\ell/d\sigma$ that, at a geodesic C_0

$$\left\{ \frac{d^2 \ell_\sigma}{d\sigma^2} \right\}_{\sigma=0} \equiv \ell_0^{-1} \int_0^1 \left\{ \frac{\partial}{\partial\sigma} \{ (v_\sigma^\lambda \nabla_\lambda v_\sigma^\alpha) g_{\alpha\beta}(\psi) h_\sigma^\beta \} d\lambda \right\}_{\sigma=0} \tag{10.5}$$

We use the fact that acting on a scalar function $\frac{\partial}{\partial\sigma} = h^\mu \nabla_\mu$ to see that at the geodesic curve C_0 10.5 reduces to, with $v = u_0$

$$\left\{ \frac{d^2 \ell_\sigma}{d\sigma^2} \right\}_{\sigma=0} \equiv \ell_0^{-1} \int_0^1 \{ h^\mu h_\alpha \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha) \}_{\sigma=0} ds \tag{10.6}$$

We compute

$$h^\mu \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha) \equiv h^\mu \nabla_\mu v^\lambda \nabla_\lambda v^\alpha + h^\mu v^\lambda \nabla_\mu \nabla_\lambda v^\alpha \tag{10.7}$$

Hence

$$h^\mu \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha) \equiv h^\mu \nabla_\mu v^\lambda \nabla_\lambda v^\alpha + h^\mu v^\lambda (\nabla_\lambda \nabla_\mu v^\alpha + R^{\cdots\cdot\alpha}_{\mu\lambda\cdots\beta} v^\beta) \tag{10.8}$$

We make appear only derivations in the direction of C by using the following condition which results from the definitions of h and v which imply that v and h commute (i.e. the Lie derivative of h along C is zero):

$$h^\mu \nabla_\mu v^\lambda = v^\mu \nabla_\mu h^\lambda. \tag{10.9}$$

Then we have

$$h^\mu \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha) \equiv v^\mu \nabla_\mu (h^\lambda \nabla_\lambda v^\alpha) + h^\mu v^\lambda R_{\mu\lambda}{}^\alpha{}_\beta v^\beta$$

and

$$h^\mu \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha) \equiv v^\mu \nabla_\mu (v^\lambda \nabla_\lambda h^\alpha) + h^\mu v^\lambda R_{\mu\lambda}{}^\alpha{}_\beta v^\beta.$$

At a geodesic, with tangent vector u, we have therefore

$$J^\alpha(h) := \{h^\mu \nabla_\mu (v^\lambda \nabla_\lambda v^\alpha)\}_{\alpha=0} = v^\mu v^\lambda \nabla_\mu \nabla_\lambda h^\alpha + h^\mu v^\lambda v^\beta R_{\mu\lambda}{}^\alpha{}_\beta. \tag{10.10}$$

We denote by $\frac{D}{D\lambda} := v^\alpha \nabla_\alpha$ the covariant derivative along C, mapping tensors defined along C into tensors of the same kind. We have found that the second variation of ℓ reads

$$\left\{ \frac{d^2 \ell(C_\alpha)}{d\alpha^2} \right\}_{\alpha=0} = \ell^{-1} \int_0^1 J^{(\alpha)}(h) h_\alpha ds \tag{10.11}$$

with

$$J^\alpha(h) \equiv \frac{D^2 h^\alpha}{D\lambda^2} - h^\mu v^\lambda v^\beta R_{\mu\lambda}{}^\alpha{}_\beta. \tag{10.12}$$

Definition 10.1 *A vector h defined along a geodesic C such that $J^\alpha(h) = 0$ is called a Jacobi field.*

Conjugate points

Definition 10.2 *Two points a and b on a geodesic C are said to be conjugate if there is a Jacobi h field on C which vanishes at a and b.*

We see on the equation of geodesic deviation that the infinitesimal distance h between nearby geodesics is a Jacobi field; the vanishing of h at a and b on C signals that other geodesics starting from a will approach again C near b, though they may not actually pass through b.

One can prove the following theorem (see Chapter 13):

Theorem 10.3 *A smooth timelike curve C connecting a to b realizes a local maximum of length between these points if and only if it is a geodesic with no conjugate point to a between a and b.*

11 Linearized Ricci and Einstein tensors

The **linearization**[13], also called **first variation**, of an operator $P : u \mapsto P(u)$ between open sets of normed vector spaces E_1 and E_2, at a point $u \in E_1$, is a linear operator acting on vectors $\delta u \in E_2$ given by the (Frechet) derivative P'_u of P at u, i.e. such that

$$\delta P := P'_u \delta u \quad \text{with} \quad P(u+\delta u) - P(u) = P'_u(u)\delta u + o(|\delta u|). \tag{11.1}$$

We consider a pseudo-Riemannian metric, with components $g_{\alpha\beta}$ in local coordinates. The relation between $g_{\alpha\beta}$ and the inverse matrix $g^{\alpha\beta}$

$$g^{\alpha\lambda} g_{\alpha\beta} = \delta^\lambda_\beta, \quad \text{the Kronecker symbol,} \tag{11.2}$$

implies that

$$\delta g^{\alpha\beta} = -g^{\alpha\lambda} g^{\beta\mu} h_{\lambda\mu}, \quad \text{where we have set} \quad h_{\lambda\mu} := \delta g_{\lambda\mu}. \tag{11.3}$$

The definition and the above relation applied to the Christoffel symbols and the Ricci tensor give by straightforward computation that $\delta\Gamma^\lambda_{\alpha\beta}$ is the following 3-tensor (indices raised with $g^{\lambda\mu}$)

$$\delta\Gamma^\lambda_{\alpha\beta} \equiv \frac{1}{2}\{\nabla_\alpha h^\lambda_\beta + \nabla_\beta h^\lambda_\alpha - \nabla^\lambda h_{\alpha\beta}\} \tag{11.4}$$

From this formula and the expression of the Ricci tensor, it follows that $\delta Ricci(g)$ is the linear operator on $h := \delta g$ given in local coordinates by

$$\delta R_{\alpha\beta} \equiv -\frac{1}{2}\nabla^\lambda \nabla_\lambda h_{\alpha\beta} + \frac{1}{2}\{\nabla_\lambda \nabla_\alpha h^\lambda_\beta + \nabla_\lambda \nabla_\beta h^\lambda_\alpha - \nabla_\alpha \nabla_\beta h^\lambda_\lambda\}$$

From this identity results the linearization of the scalar curvature

$$\delta R \equiv g^{\alpha\beta}\delta R_{\alpha\beta} + R_{\alpha\beta}\delta g^{\alpha\beta},$$

where $g^{\alpha\beta}\delta R_{\alpha\beta}$ is the divergence of a vector, namely:

$$g^{\alpha\beta}\delta R_{\alpha\beta} \equiv -\nabla_\lambda\{\nabla^\lambda h^\alpha_\alpha - \nabla_\alpha h^{\lambda\alpha}\}. \tag{11.5}$$

12 Second derivative of the Ricci tensor

The second (Fréchet) derivative, also called **second variation**, of the operator P at $u \in E_1$ is, if E_1 is a Banach algebra, a quadratic form $P"_{u^2}$ in δu such that

$$\delta^2 P := P"_{u^2}(\delta u, \delta u) \quad \text{with} \tag{12.1}$$

$$P(u+\delta u) - P(u) = P'_u(u)\delta u + \frac{1}{2}P"_{u^2}(u)(\delta u, \delta u) + o(|\delta u|^2). \tag{12.2}$$

[13] See for instance CB-DM1, II A, in particular Problem 1.

Straightforward though lengthy computation gives[14] for the **second variation** of the Ricci tensor at a metric g a quadratic form in $h := \delta g$:

$$\delta^2 R_{\alpha\beta} := R"_{\alpha\beta}(g)(h,h). \tag{12.3}$$

which reads

$$\begin{aligned}\delta^2 R_{\alpha\beta} \equiv & -h^{\lambda\mu}\{\nabla_\lambda(\nabla_\alpha h_{\beta\mu} + \nabla_\beta h_{\alpha\mu} - \nabla_\mu h_{\alpha\beta}) - \nabla_\alpha\nabla_\beta h_{\lambda\mu}\} \\ & - \nabla_\lambda h^{\lambda\mu}(\nabla_\alpha h_{\beta\mu} + \nabla_\beta h_{\alpha\mu} - \nabla_\mu h_{\alpha\beta}) + \frac{1}{2}\nabla_\beta h^{\lambda\mu}\nabla_\alpha h_{\lambda\mu} \\ & + \frac{1}{2}\nabla^\lambda h^\rho_\rho(\nabla_\alpha h_{\beta\lambda} + \nabla_\beta h_{\alpha\lambda} - \nabla_\lambda h_{\alpha\beta}) + \nabla_\lambda h^\mu_\alpha \nabla^\lambda h_{\beta\mu} - \nabla_\lambda h^\mu_\alpha \nabla_\mu h^\lambda_\beta.\end{aligned} \tag{12.4}$$

[14] Choquet-Bruhat, Y. (2000) *Ann. Phys. (Leipzig)* **9**(3–5), 258–66.

II

SPECIAL RELATIVITY

1 Newton's mechanics

1.1 *The Galileo–Newton spacetime*

In the spacetime of Galileo and Newton the simultaneity of two events is a notion independent of the observers. Space and time are absolute objects[1] which exist independently of matter and events which may happen in them. The mathematical model of Newtonian spacetime is the direct product of R, where time varies, and space. The space itself is taken to be[2] a Euclidean space E_3, that is R^3 with the Euclidean metric, which reads in standard (called Cartesian) coordinates:

$$ds^2 = \sum_{i=1,2,3} (dx^i)^2 \tag{1.1}$$

To link this mathematical model with observations, one has to identify points of this Euclidean space with observed objects. This identification was made with better and better approximation through objects first at rest with respect to the Earth, then at rest with respect to a Cartesian frame centred at the Sun and with axes directed towards specific stars[3]. Finally the physicists of the previous century introduced a mysterious medium, called aether, invisible and intangible, though with all the properties of a solid.

Measurements. The isometry group of the Euclidean model of space predicts the existence of solid bodies undergoing free motions. Measures can be done by comparison with a standard of length, defined by a piece of metal deposited in Sèvres. For distances which are too large to be compared directly with a copy of the standard metre one can use the properties of Euclidean geometry, such as triangulation. The existence of clocks measuring[4] the absolute time is predicted by the periodic phenomena resulting from Newtonian dynamics.

[1] Recall that in the same years Leibnitz was in disagreement with this postulate.

[2] Inspired by the Galilean invariance, Trautman (Trautman, A. (1964) In *Brandeis Summer Institute in Theoretical Physics*, Prentice Hall) prefers to say that Newtonian spacetime is a fibre bundle with base the time line R and fiber R^3. Such a spacetime is a nice mathematical model, but physics needs to identify the representative of a fibre with observable objects.

[3] Remark that for the physical identification of absolute space, Newtonian physics anticipates Mach's principle, for which space is determined by the bulk of surrounding matter.

[4] At least approximately, for instance the small oscillations of a pendulum.

1.2 *Newton's dynamics – the Galileo group*

The fundamental law of Newtonian dynamics for a particle supposed to be pointlike is

$$F = m\gamma, \tag{1.2}$$

where m is a constant, the mass of the particle, also called inertial mass, γ is its acceleration in absolute time with respect to absolute space, and F is a phenomenological vector, the applied force. The components of γ in an inertial frame, i.e. in Cartesian coordinates x^i for the particle in the absolute space E_3, are the second partial derivatives $\gamma^i = \frac{d^2x^i}{dt^2}$.

Newton's law (1.2) is invariant under the following time-dependent change of coordinates in Newton's absolute space E^3

$$x'^i = x^i + v^i t + x'^i_0 \tag{1.3}$$

where the $v^i s$, x'^i_0 are constants. The corresponding t-dependent Cartesian frame in E^3 is in uniform translation with respect to the original absolute space. Such reference frames are called **inertial** frames. The set of transformations (1.3) forms a group, called the **Galileo group**. It was remarked already by Galileo that a uniform in time translation of a boat cannot be detected by observers in the hold: more generally, all the physical laws of Newton's mechanics were supposed to be invariant under the Galileo group, in the sense that they admit the same formulation in all inertial systems.

2 Maxwell's equations

Maxwell's equations are the great success of 19th century theoretical physics, and are still valid today. They unify the various physical laws which govern the electric and magnetic fields (E, H). These equations in non-inductive media, with proper choice of units ("geometric" units; see Section 2.6), read in 3-vector notation on R^3:

$$curl\ E = -\frac{\partial H}{\partial t} \quad \text{(Faraday's law)} \tag{2.1}$$

$$curl\ H = j + \frac{\partial E}{\partial t} \quad \text{(Ampère–Maxwell's law)} \tag{2.2}$$

$$Div\ E = q \quad \text{(Coulomb's law)} \tag{2.3}$$

$$Div\ H = 0 \quad \text{(Gauss's law, non-existence of magnetic charges)} \tag{2.4}$$

The scalar q is the charge density. The vector j is the electric current.

Equations (2.1) and (2.2) are evolution equations; (2.3) and (2.4) are "constraints", i.e. do not contain time derivatives.

In the absence of electric current and charge this set of equations implies that E and H satisfy the following wave equations (Δ is the Laplace operator of

Euclidean 3-space):

$$-\frac{\partial^2 H}{\partial t^2} + \Delta H = 0, \quad -\frac{\partial^2 E}{\partial t^2} + \Delta E = 0. \tag{2.5}$$

The various laws written above were interpreted before Einstein with t the absolute time and E_3 the absolute space determined by the mysterious medium aether. Maxwell's equations implied then that the electric and magnetic fields, and also light[5], which is an electromagnetic phenomenon, propagate in vacuum with velocity equal to 1, independent of the time t and the location in this absolute space.

The problem for Newton's mechanics is that Maxwell's equations are not invariant under the Galileo group. Indeed the transformation (1.3) leaves invariant the vectors E and H on R^3 and their curls, but not their time derivatives. It holds that

$$E'(x') = E(x' - vt), \quad \frac{\partial E'(x')}{\partial t} = \left(\frac{\partial E}{\partial t} - v^i \frac{\partial E}{\partial x^i}\right)(x' - vt). \tag{2.6}$$

It was the mathematical genius of Poincaré, and Lorentz, to discover the group which leaves invariant Maxwell's equations, but neither of them discarded Newton's absolute time, nor the aether filling the absolute space. The Poincaré group is in fact the group of isometries of a flat Lorentzian 4-manifold, the Minkowski spacetime M_4.

3 Minkowski spacetime

3.1 *Definition*

Special Relativity, discovered by Einstein in 1905, replaces the Newtonian absolute space E_3 and absolute time belonging to R by the Minkowski spacetime[6] M_4. It is the $3+1$-dimensional manifold $R^3 \times R$ endowed with a Lorentzian, flat metric. More generally one defines the Minkowski spacetime M_{n+1} as the manifold R^{n+1} endowed with a metric which reads[7] in canonical, called **inertial**, coordinates:

$$\eta = -(dx^0)^2 + \sum_{i=1,\dots,n} (dx^i)^2. \tag{3.1}$$

In inertial coordinates the covariant derivative ∇_α in the Minkowski metric coincides with the partial derivative $\frac{\partial}{\partial x^\alpha}$.

We have defined in Chapter 1 general Lorentzian manifolds and the corresponding notions of causality.

[5] It was known before that the speed of light is finite, measured by Römer in 1675 by the observation of the eclipses of Jupiter's satellites.

[6] Minkowski (1908).

[7] We have adopted the MTW (Misner, Thorne, Wheeler) signature convention, though many authors still use the signature $(+---)$, which is more convenient in some problems. The choice is irrelevant for mathematics and physics, but must be kept in mind.

Remark 3.1 *The Minkowski spacetime has a vector space and an affine structure compatible with its metric. It leads often to the identification of this spacetime with its tangent bundle, a fact which does not hold for a general manifold.*

3.2 *Maxwell's equations on M_4*

Theorem 3.2 *Maxwell's equations (2.1–4) can be written as a pair of equations for an exterior 2-form F on Minkowski spacetime M_4; this 2-form represents both the electric and magnetic fields. The equations are*

$$dF = 0 \tag{3.2}$$

where d is the exterior differential; equivalently, in arbitrary coordinates

$$\nabla_\alpha F_{\beta\gamma} + \nabla_\gamma F_{\alpha\beta} + \nabla_\beta F_{\gamma\alpha} = 0 \tag{3.3}$$

and

$$\nabla \cdot F = J; \tag{3.4}$$

that is, in arbitrary coordinates with indices raised with the Minkowski metric

$$\nabla_\alpha F^{\alpha\beta} = J^\beta, \quad \textit{with} \quad J^i = -j^i, \quad J^0 = q. \tag{3.5}$$

Proof Consider the 2-form on the 4-dimensional manifold $R^3 \times R$:

$$F \equiv \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta \tag{3.6}$$

with components in inertial coordinates $(x^\alpha) = (x^i, x^0 = t)$ given by:

$$F_{i0} = E_i, \quad F_{23} = H_1, \quad F_{31} = H_2, \quad F_{12} = H_3. \tag{3.7}$$

It is straightforward to check that Equations (2.1) and (2.4) are equivalent to

$$dF = 0$$

The proof that Equations (2.2) and (2.3) are equivalent to $\nabla.F = J$ is given in CB-DM1, Chapter 5, Problem 1. It is easy to check this equation in its inertial coordinates form. □

The vector J is the 4-dimensional electric current. Equation (3.4) implies its conservation, that is

$$\nabla_\alpha J^\alpha = 0. \tag{3.8}$$

This equation could have been deduced directly from Equations (3.3) and (3.5) in inertial coordinates.

Since the operator d does not depend on the spacetime metric, and $\nabla.F$ is invariant under the group leaving the metric invariant, Maxwell's equations are invariant under the isometry group of Minkowski spacetime. However, under such an isometry the space and time components of the electromagnetic field will be mixed. Vectors observed as electric and magnetic fields depend on the corresponding inertial observers. This fact has been checked over and over in laboratories.

Maxwell tensor

The Maxwell tensor τ is the symmetric 2-tensor given on Minkowski spacetime by

$$\tau_{\alpha\beta} := F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4}\eta_{\alpha\beta}F^{\lambda\mu}F_{\lambda\mu}. \tag{3.9}$$

The definitions (3.7) of E and H show that the components $\tau_{0\alpha}$ of τ in inertial coordinates read as follows, with ε_{ijl} totally antisymmetric and $\varepsilon_{123} = 1$:

$$\tau_{00} = \frac{1}{2}(E^2 + H^2), \quad \tau_{0i} = -\varepsilon_{ijl}E^j H^l; \tag{3.10}$$

The component $\tau_{00} = \tau^{00}$ is the energy density of the electromagnetic field, and $\tau^{0i} = -\tau_{0i}$, $i = 1, 2, 3$, are the components of the Poynting energy flux vector P, given in vector product notation on R^3 by

$$P = E \wedge H. \tag{3.11}$$

Lemma 3.3 *Modulo Maxwell's equations it holds that:*

$$\nabla_\alpha \tau^{\alpha\beta} = J^\lambda F_{\beta\lambda}. \tag{3.12}$$

Proof Straightforward calculation gives

$$\nabla_\alpha \tau^\alpha_\beta \equiv (\nabla_\alpha F^{\alpha\lambda})F_{\beta\lambda} + F^{\alpha\lambda}\nabla_\alpha F_{\beta\lambda} - \frac{1}{2}F^{\lambda\mu}\nabla_\beta F_{\lambda\mu}.$$

Hence, changing names of indices and using antisymmetries,

$$\nabla_\alpha \tau^\alpha_\beta = (\nabla_\alpha F^{\alpha\lambda})F_{\beta\lambda} + \frac{1}{2}F^{\alpha\lambda}(\nabla_\alpha F_{\beta\lambda} + \nabla_\lambda F_{\alpha\beta} + \nabla_\beta F_{\lambda\alpha}). \tag{3.13}$$

The Equations (3.3) and (3.5) imply the result. □

The vector $J^\lambda F_{\beta\lambda}$ is called the **Lorentz force**.

Equations (3.12) written in inertial coordinates are, when the Lorentz force is identically zero, the usual equations of conservation of energy and momentum.

Remark 3.4 *Maxwell's equations (3.2) and (3.4) and the Maxwell tensor can be written for a 2-form on a manifold of arbitrary dimension; the identity (3.13) is still valid. The particular form (2.1–4) holds only in space dimension 3. The magnetic vector H is then the Hodge adjoint of the 2-form on 3-space induced by F.*

4 Poincaré group

The isometry group of the Minkowski metric η is called the **Poincaré group**. Equivalently it is the group of global coordinate changes on R^{3+1} which preserve the form (3.1) of η; that is the group of changes of inertial frames.

Since the mathematics are the same and since that, in modern times, higher dimensional theories are considered in view of the unification of fundamental fields, we will work with the general case of an $n+1$-dimensional Minkowski spacetime. Its Poincaré group is the group of diffeomorphisms of R^{n+1} given by the n invertible functions

$$x^\alpha = f^\alpha(x'^1, \ldots, x'^n) \tag{4.1}$$

such that

$$-(dx^0)^2 + \sum_{i=1,\ldots,n} (dx^i)^2 = -(dx'^0)^2 + \sum_{i=1,\ldots,n} (dx'^i)^2. \tag{4.2}$$

Translations in R^{n+1}, $x^\alpha = x'^\alpha + c^\alpha$, are obviously a subgroup of the Poincaré group. It is a transitive group on R^{n+1}. The Poincaré group is the semi-direct product of R^{n+1} by the isotropy group, which is the same at different points. This group, called the **Lorentz group**, is isomorphic with the group of matrices $L \equiv (L^\alpha_{\alpha'})$ leaving invariant the following quadratic form in the vector space R^{n+1}:

$$-(v^0)^2 + \sum_{i=1,\ldots,n} (v^i)^2. \tag{4.3}$$

That is, it holds that for each set $\{v^{\alpha'}\}$:

$$-(L^0_{\alpha'} v^{\alpha'})^2 + \sum_{i=1,\ldots,n} (L^i_{\alpha'} v^{\alpha'})^2 \equiv -(v'^0)^2 + \sum_{i=1,\ldots,n} (v'^i)^2. \tag{4.4}$$

5 Lorentz group

5.1 *General formulae*

Equation (4.4) with $v^{0'} = 1$, $v^{i'} = 0$ implies that

$$(L^0_{0'})^2 = 1 + \sum_{i=1,\ldots,n} (L^i_{0'})^2 \geq 1. \tag{5.1}$$

The **orthochronous** Lorentz group is the subgroup which preserves the orientation of R defined by the coordinate x^0; i.e. is such that:

$$L^0_{0'} > 0 \tag{5.2}$$

The elements with $L^0_{0'} < 0$ reverse the orientation of R. They do not constitute a group.

Equation (4.4) implies that the determinant of L is equal to $+1$ or -1. **Proper** Lorentz transformations are those which preserve the spacetime orientation; i.e. are such that

$$Det(L) = 1. \tag{5.3}$$

Standard geometrical considerations show that every proper Lorentz transformation can be written

$$L = R_1 L_S R_2 \tag{5.4}$$

where R_1 and R_2 are elements of the rotation group $O(n)$ and L_S is a so-called **special Lorentz** transformation, acting only in the time and one space direction; namely such that Equations (5.1) reduce to

$$(L^0_{0'})^2 - (L^1_{0'})^2 = 1, \tag{5.5}$$

$$(L^1_{1'})^2 - (L^0_{1'})^2 = 1, \tag{5.6}$$

$$L^0_{0'} L^0_{1'} - L^1_{1'} L^1_{0'} = 0. \tag{5.7}$$

The general solution of these equations is, for an orthochronous transformation:

$$L^0_{0'} = L^1_{1'} = \cosh\varphi, \quad L^1_{0'} = L^0_{1'} = \sinh\varphi. \tag{5.8}$$

We extend to the whole of M_{n+1}, by using its affine structure, the isometry given by the special Lorentz transformation. We set $x^0 = t$ and $V = \tanh\varphi$ and remark that $V \leq 1$. We obtain the transformation law

$$t - \tau = \frac{t' - \tau' + V(x'^1 - \xi'^1)}{\sqrt{1 - V^2}}, \quad x^1 - \xi^1 = \frac{x'^1 - \xi'^1 + V(t' - \tau')}{\sqrt{1 - V^2}}, \tag{5.9}$$

where (x^1, t) and (ξ^1, τ) are the coordinates of two points in the (x^1, t) plane in an inertial system, and the prime quantities are the coordinates of these points in another inertial system in uniform translation along the x^1 axis with respect to the first one. The relative velocity of this translation is dx'^1/dt for fixed x^1, while the relative velocity of the unprimed frame with respect to the primed one is dx^1/dt' for fixed x'^1. These velocities resulting from the above formulas are found to be:

$$\frac{dx^1}{dt'}|_{x'^1 = const} = \frac{V}{\sqrt{1 - V^2}} \quad \text{and} \quad \frac{dx'^1}{dt}|_{x^1 = const} = \frac{-V}{\sqrt{1 - V^2}}. \tag{5.10}$$

5.2 *Transformation of electric and magnetic vector fields (case $n = 3$)*

The transformation of an electromagnetic field F under a Lorentz transformation is:

$$F_{\alpha\beta} \equiv \frac{\partial x^{\alpha'}}{\partial x^\alpha} \frac{\partial x^{\beta'}}{\partial x^\beta} F_{\alpha'\beta'} = L^{\alpha'}_\alpha L^{\beta'}_\beta F_{\alpha'\beta'}. \tag{5.11}$$

In the case $n = 3$, for a special Lorentz transformation, this gives

$$E = E' + v \wedge H, \quad H = H' - v \wedge E. \tag{5.12}$$

where v is now the vector supported by the x^1 axis with length $V/\sqrt{1 - V^2}$. The formulae (5.12) are currently used, even for non-relativistic charged fluids.

5.3 *Lorentz contraction and dilatation*

The results obtained by Lorentz led him and Poincaré to controversial studies on the dynamics of charged bodies, before the new physics introduced by Einstein. We give some of their results.

- Consider two events simultaneous in the primed frame: $t' = \tau'$. The formulas (5.11) give

$$x^1 - \xi^1 = \frac{x'^1 - \xi'^1}{\sqrt{1 - V^2}}, \tag{5.13}$$

  Hence the spatial distance observed in the primed frame is smaller than the one observed in the unprimed one. This is the **Lorentz contraction**.
- Consider two events with the same space location in the primed frame: $x'^1 = \xi'^1$. Then

$$t - \tau \geq t' - \tau', \tag{5.14}$$

  This is the **Lorentz dilatation**.

The Lorentz contraction and dilatation are not intrinsic phenomena. They are relative to the observers linked with the reference frames, as is obvious from the fact that they are reversed by exchanging the roles of these frames when defining simultaneity or space coincidence.

6 Special Relativity

6.1 *Proper time*

Lorentz and Poincaré had made mathematical studies, but it was Einstein[8] (1905) who made the conceptual jump to discard Newton's absolute time and space as physical realities. Einstein, in Special Relativity, replaces the direct product $E_3 \times R$ by the Minkowski spacetime M_4. An event is now a point in spacetime. Its history is a timelike trajectory in M_4. One could consider as simultaneous a pair of events located on a spacelike 3-manifold, but this is a very non-unique definition. Simultaneity, and hence space, is a relative notion. The Einsteinian revolution was to discover that time is also relative. The parameter t appearing in the Minkowski metric has no physical meaning: the quantity measurable by well-defined clocks (mechanical, atomic or biological) is the length of their timelike trajectories in spacetime, called the **proper time**. More precisely, if $C : t \mapsto C(t)$, $t \in [t_1, t_2]$ is a future causal curve parametrized by t joining two points of M_4 the proper time along that curve is the parameter-independent quantity

$$\int_{t_1}^{t_2} \left\{ -\eta \left(\frac{dC}{dt}, \frac{dC}{dt} \right) \right\}^{\frac{1}{2}} dt. \tag{6.1}$$

[8] See the historical discussion in Damour, T. (2005) *Si Einstein m'était conté*, Cherche Midi, translated by E. Novak (2006) as *Once Upon Einstein*, A. K. Peters Ltd, Wellesley MA.

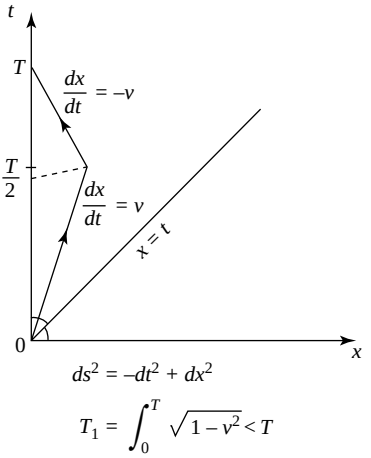


Figure 2.1 The twin paradox

Relativity postulates the existence of universal clocks defined by a specific physical phenomenon which measure the proper time. Such clocks are obtained by using the frequency of emission of specific radiations by atoms, predicted by quantum theory to have a constant universal value. The actually adopted standard clock is the caesium atom which presents a particularly stable – of the order of 10^{-13} – transition between two particular energy levels. The second is now defined through the time measured by the caesium clock.

The discovery by Einstein of the proper time has been proved accurate in all experiments.

In a Lorentzian manifold the length of timelike curves joining two points has a local maximum for a timelike geodesic (see Chapter 1). Geodesics of Minkowski spacetime are represented by straight lines in inertial coordinates. In particular, the line where only the parameter t of inertial coordinates varies is a timelike geodesic. Therefore the time measured by an observer at rest in some inertial coordinate system is greater than the time measured by a traveller who does not follow such a straight line between their separation and their reunion. This so-called "**twin paradox**" (see Fig. 2.1)[9], has been long verified on elementary particles in modern accelerators. The reality of proper time has also been checked recently in an airplane flight of a caesium clock which was late with respect to

[9] There is no paradox because the two twins do not have the same history: one describes a geodesic in Minkowski spacetime, and the other does not; in fact he has to use a motor to follow his trajectory. This physical effect is not to be confused with the apparent time dilatation, which is a reciprocal effect.

a similar clock that remained on ground[10]. In long space journeys this could be verified with the human biological clock.

Remark 6.1 *There are well-defined clocks to measure time (proper time) in Relativity, but there is no standard of length because of the lack of an absolute notion of simultaneity.*

Remark 6.2 *Special Relativity says that, gravitation being neglected, the arena of noticeable phenomena is to be identified with a Minkowski spacetime. It supposes that it is possible to make this identification observationally; in particular to physically determine a priviledged Minkowskian inertial reference frame. It seems that such a frame is in general empirically chosen, depending on the circumstances, as in the Newton–Galileo case. The identification of physical space with some Euclidean space E_3 is made by consideration of the bulk of corresponding matter. The timelines of observers at rest in that space are identified as orthogonal to space in the Minkowski metric.*

6.2 *Proper frame and relative velocities*

A velocity, even in Galileo–Newton mechanics, is always defined with respect to some observer. The problem is more complex in Relativity since there is neither absolute time nor absolute simultaneity. In Relativity an observer is defined by its world line, a timelike curve. The norm (in the Minkowski metric) of its timelike tangent vector is dependent on the choice of the parameter on the curve; this norm has no physical meaning if the parameter is not specified. One labels as a "unit tangent vector" u this tangent vector normalized by

$$\eta(u,u) = -1. \tag{6.2}$$

The **proper frame** of the observer (not necessarily following a geodesic) at some point of spacetime, is an orthonormal Lorentzian frame with u as its timelike axis. Consider an object following a causal line with tangent v at the considered point. If in the proper frame the vector v has time component v^0 and space components v^i, one says[11] that the object has velocity V with respect to this observer, where V is a space vector[12] with components

$$V^i = \frac{v^i}{v^0}. \tag{6.3}$$

[10] The effect of gravity on proper time will be considered in the following chapter. Here it is the observer on the ground which is (approximately) inertial, while the airplane, using its motor, is not.

[11] This definition comes to associate with the observer at some point of spacetime the inertial system in which he is momentarily at rest.

[12] That is, orthogonal to u.

Since v is causal, it holds that

$$|V| \equiv \left\{ \sum_i (V^i)^2 \right\}^{\frac{1}{2}} \leq 1, \tag{6.4}$$

with $|V| = 1$, the speed of light, if and only if v is a null vector.

In particular, consider two observers at a point of spacetime with unit velocities u and u'. Choose their proper frames such that they are linked by a special Lorentz transformation, i.e. $e_2 = e'_2$, $e_3 = e'_3$. The velocity of the primed observer with respect to the unprimed observer is then along the axis e_1 and given by (with $\tilde{u}^\alpha$ components of u' in the unprimed frame)

$$V = \frac{\tilde{u}^1}{\tilde{u}^0} = \frac{L^1_{0'}}{L^0_{0'}} \tag{6.5}$$

Since the components of u' in the primed frame are $u'^1 = 0$, $u'^0 = 1$.

Addition of velocities

As foreseen from the overturning of the notion of absolute time and simultaneity, relativistic relative velocities do not add simply as in Newtonian mechanics.

Consider two systems of inertial coordinates (t, x^1, x^a) and (t', x'^1, x^a). They define at each point of spacetime two Lorentz frames, linked by a special Lorentz transformation, in relative motion with velocity V. The velocity of a given point particle is dx^1/dt with respect to the unprimed frame and dx'^1/dt' with respect to the primed frame. The formula (5.9) implies that

$$\frac{dx^1}{dt} = \frac{dx'^1 + V dt'}{dt' + V dx'^1}, \tag{6.6}$$

Therefore, setting $\frac{dx^1}{dt} = U$, $\frac{dx'^1}{dt'} = U'$,

$$U = \frac{U' + V}{1 + U'V} \tag{6.7}$$

This coincides to a first approximation with the classical formula of addition of velocities when UV' is small with respect to 1, the light velocity.

Formula (6.7) implies that $U = 1$ when $U' = 1$: the speed of light is independent of the relative velocity V of different observers. This property, very surprising in the framework of the Galilean kinematics, but already contained in Maxwell's equations, was found experimentally by Michelson[13], and verified more accurately by Morley. **Michelson and Morley** measured the speed of light in the direction of the orbital velocity of the Earth and in a transversal direction. To the surprise of the scientific world they found the same number in both cases, to an accuracy higher than possible experimental error. Since then, it has been

[13] It was known before that the speed of light is finite, measured by Römer in 1675 by the observation of the eclipses of Jupiter's satellites.

verified with greater and greater accuracy, in a medium which does not interfere with this velocity, vacuum. The most recent experiments give, in the usual units, $2.99792458 \times 10^8\,\mathrm{m\,s^{-1}}$. The Michelson–Morley experiment was in part[14] a source of inspiration to Einstein to discover Special Relativity.

Since the speed of light in vacuum is a universal constant, it can be used to define the standards of length and time from one another. Physics has shown that time can be measured more accurately (with a caesium clock) than length (as it was previously with a wavelength of krypton). It has been decided by scientific authorities that the metre is now defined to be the distance covered by light in $(2.99792458)^{-1} \times 10^{-8}$ seconds. The speed of light is therefore fixed to be a universal constant. Unless otherwise specified, we follow the mathematical usage to choose units of length and time such that this constant is equal to 1 (this is its "geometric" value, which we have taken in writing the Maxwell equations). In comparison with observations or experiments, other units may be more appropriate.

7 Dynamics of a pointlike mass

7.1 *Newtonian law*

The Newtonian equation of motion (1.2) of a particle with mass m subjected to the force f can be written

$$\frac{d(mv)}{dt} = f. \tag{7.1}$$

with v its velocity with respect to a Galilean inertial frame and t the Newtonian absolute time. When $f = 0$, the particle is in uniform rectilinear motion in all Galilean inertial frames.

From (7.1) one obtains the Newtonian energy equation, where a dot denotes the Euclidean scalar product:

$$\frac{d}{dt}\left(\frac{1}{2}mv^2\right) = f \cdot v. \tag{7.2}$$

If one wants to write Newton's law (7.1) in a non-Galilean frame, one must add to f the so called inertial forces due to the motion of the considered frame with respect to a Galilean one.

7.2 *Relativistic law*

The trajectory of a pointlike massive particle in Minkowski spacetime is a timelike curve. Since there is no absolute time to define its velocity, one consider its **unit velocity** u, the tangent vector to its trajectory parametrized by proper

[14] It seems that Einstein was more inspired by the Faraday induction laws than by the Michelson–Morley experiment.

time (the Minkowskian arc length s). The components of u are, in arbitrary coordinates:

$$u^\alpha = \frac{dx^\alpha}{ds}, \quad \text{they satisfy} \quad u^\alpha u_\alpha = -1. \tag{7.3}$$

The acceleration of the particle is the derivative of u in the direction of itself, $u^\alpha \nabla_\alpha u$. The time-dependent, coordinate-dependent, Newtonian equations on space (7.1) and (7.2) are replaced in relativistic dynamics by a spacetime, coordinate-independent, equation

$$u^\alpha \nabla_\alpha (m_0 u^\beta) = F^\beta, \tag{7.4}$$

where m_0 is a scalar, called the rest mass of the particle, and F is now a spacetime vector, called the 4-force. Using $u_\beta \nabla_\alpha u^\beta = 0$, a consequence of $u_\beta u^\beta =$ constant, (7.4) gives

$$-u^\alpha \partial_\alpha m_0 = u_\beta F^\beta; \tag{7.5}$$

Therefore m_0 is a constant[15] on the trajectory if and only if $u_\beta F^\beta = 0$; i.e. the 4-force F is orthogonal (in the Minkowski metric) to the trajectory. If this is not the case and the spacetime vector F is given, the quantity m_0 is another unknown for Equations (7.4).

A physical quantity which can be defined also for particles with zero rest mass is the **energy momentum**, a vector tangent to the trajectory and such that

$$-\eta(p,p) \equiv -p_\alpha p^\alpha = (m_0)^2. \tag{7.6}$$

Equation (7.4) reads now:

$$u^\alpha \nabla_\alpha p^\beta = F^\beta. \tag{7.7}$$

The component p^0 in some Lorentzian frame is the energy of the particle with respect to this frame. The components p^i define its momentum. This splitting is frame dependent. In the rest frame of a massive[16] particle it holds that $u^i = 0$ and $u^0 = 1$; hence in that frame

$$p^0 = m_0. \tag{7.8}$$

Remark 7.1 *If m_0 is a constant Equation (7.4) reads (see I.7.4)*

$$m_0 \left\{ \frac{d^2 x^\beta}{ds^2} + \Gamma^\beta_{\alpha\lambda} \frac{dx^\alpha}{ds} \frac{dx^\lambda}{ds} \right\} = F^\beta. \tag{7.9}$$

In an inertial reference frame the connection coefficients vanish and Equation (7.9) reduces to an analogue of Newton's equation in a Galilean frame. This remark leads to interpretation of the term in Γ as a kind of inertial force.

[15] The rest mass of a molecule is constant as long as its structure is not changed by chemical reactions. The rest mass of an atom is modified by its absorption or emission of photons.

[16] Particles with zero mass have no rest frame.

7.3 *Equivalence of mass and energy*

Newtonian interpretation of the relativistic equation

We denote by U the relative velocity of the particle with respect to some Minkowskian inertial coordinates where its unit velocity is u; that is[17], since on the trajectory $\frac{dx^0}{ds} \equiv \frac{dt}{ds} =: u^0$

$$U^i := \frac{dx^i}{dt} = \frac{dx^i}{ds}\frac{ds}{dt} = \frac{u^i}{u^0}; \tag{7.10}$$

hence:

$$u^i = \frac{U^i}{\sqrt{1-|U|^2}}, \qquad u^0 = \frac{1}{\sqrt{1-|U|^2}}, \tag{7.11}$$

with

$$|U|^2 = \sum_i (U^i)^2. \tag{7.12}$$

Using previous relations we see that Equations (7.4) with index i read

$$\frac{d}{dt}\frac{m_0 U^i}{\sqrt{1-|U|^2}} = F^i\sqrt{1-|U|^2}. \tag{7.13}$$

This formula coincides with the Newtonian formula if one sets

$$m = \frac{m_0}{\sqrt{1-|U|^2}}, \quad \text{and} \quad F^i\sqrt{1-|U|^2} = f^i. \tag{7.14}$$

The relation (7.5) implies then $F^0 = f.U$, and Equation (7.4) with index zero becomes

$$\frac{dm}{dt} = F^0\sqrt{1-|U|^2} = f.U. \tag{7.15}$$

If we interpret the relativistic formulae in a Newtonian context we see that the Newtonian mass m is a time-dependent quantity, even for a constant rest mass m_0. The variation of m is linked with the work of the force f.

For $|U|$ small with respect to 1, we have $\sqrt{1-|U|^2} \cong 1 + \frac{1}{2}|U|^2$. Then

$$m \cong m_0\left(1 + \frac{1}{2}|U|^2\right). \tag{7.16}$$

Therefore, in this approximation, m is the sum of the rest mass m_0 of the particle and its kinetic energy.

[17] The formula below agrees with the definition given in Section 6.4, with V now denoted u.

Equivalence of mass and energy

Equation (7.4), as well as the addition of the "kinetic energy" to the rest mass m_0 in the expression for m in the approximation of small velocities led Einstein to his famous postulate of the equivalence of mass and energy[18], with a conversion factor of the order of c^2, c being the speed of light, taken to be 1 in the expression used for the Minkowski metric (geometrized units). The equivalence of mass and energy has been verified in nuclear reactions in spectacular fashion. The energy can be created by the fission of a uranium atom which is induced by a collision with a neutron. The sum of the rest masses of the incoming particles is greater than that of the post-fission particles, the difference being seen as radiated energy. Nuclear fusion of particles into a single particle with rest mass smaller than the sum of the rest masses of the incoming particles also produces energy in accord with $E = mc^2$. The fusion of two hydrogen atoms into one helium atom is the main source of the heat dispensed by our Sun. Researchers are actively trying to reproduce it on Earth. Constant rest masses are assigned to elementary particles, photon (rest mass zero), electron, proton, neutrino[19],.... The proton has now been experimentally decomposed into its constituent quarks.

8 Continuous matter

Our never-ending improvement in the exploration of reality has made us see that matter is discontinuous at all scales: galaxies in the cosmos, stars in galaxies, molecules in stars, "elementary" particles in molecules, strings,.... But at some scales, which we call "macroscopic", we are not interested in the impossible task[20] of following the individual motions of constituents. We wish to describe the behaviour of volumes, small at the observed scale, but large at the scale of the constituents, which we call particles, of the model that we are studying. Particles eventually go in and out of such a volume. If it is possible to define, pointwise on spacetime, measurable macroscopic quantities which characterize at the interesting scale the behaviour of matter, we call it continuous matter. Real matter is much too complicated to be represented by a single model. In fact, all models are only an approximation of a type of matter. Roughly speaking, fluids are matter models where in the absence of external forces defined on spacetime the only action on an elementary volume is an interface action with neighbour elements, which do not prevent them from slipping against each other. Perfect fluids are those for which the antislipping force has minimal action, being always orthogonal to the interface.

[18] Possibly all ultimate elementary particles have zero rest mass and the positive rest mass of the particles which appear to us as elementary is only an interaction energy.

[19] After a long time when the rest mass of the neutrino was believed to be zero, it is now established that this rest mass is very small, but not zero.

[20] Without speaking of the intrusion of quantum mechanics!

We will not in this chapter deduce the laws of fluid dynamics from the dynamics of pointlike particles[21]. This will be done directly in General Relativity for rarefied gases in the context of kinetic theory.

In Newtonian mechanics with absolute time and simultaneity, the state of a fluid is characterized, if one leaves aside thermodynamics considerations, by its density function ρ and its flow vector field v. These are absolute time-dependent quantities on Euclidean space E_3. They satisfy on the one hand the conservation of matter equation

$$\frac{\partial \rho}{\partial t} + \partial_i(\rho v^i) = 0, \tag{8.1}$$

and on the other hand the equations of motion obtained by passing to the pointlike limit of the Newtonian dynamical laws applied to a small volume in E_3. These equations are

$$\frac{d(\rho v^i)}{dt} + \partial_k t^{ik} = 0, \tag{8.2}$$

where t^{ik} is the stress tensor, deduced in the case of fluids from the force X which the surrounding matter imposes at a point to an element of surface with normal n by the linear relation

$$X^i = t^{ij} n_j. \tag{8.3}$$

The tensor t depends on the nature of matter. It can be proved to be symmetric when the matter has no intrinsic momentum density.

In relativistic dynamics the fundamental macroscopic, spacetime quantities characterizing a fluid are the flow vector field u, a unit vector[22], and the energy density function μ. A Lorentzian frame with timelike vector u is called the **comoving** or **proper frame** of the fluid. The function μ on spacetime is the time component in the proper frame of the energy momentum vector field $p = \mu u$.

If there are no interactions (chemical, nuclear, non-elastic shocks, ...) modifying the nature of the underlying particles, one can define for the fluid a proper mass density r, also called the particle number density. This scalar function on spacetime satisfies a conservation[23] equation which reads in arbitrary coordinates

$$\nabla_\alpha(r u^\alpha) = 0. \tag{8.4}$$

8.1 *Case of dust (incoherent matter)*

The dust model of matter is a good approximation to other models, due to the high conversion factor in usual units between rest mass and other types of energy.

[21] It is still more difficult in Einsteinian than in Newtonian dynamics, due to the fact that the proper times of distinct particles hidden in continuous matter are in general different.

[22] Except for flows of particles with zero rest mass, where u is a null vector.

[23] See the justification of this name in Chapter 9.

A fluid is called dust if neighbouring volume elements exert no action on each other. It is then supposed that the flow lines are geodesics; that is,

$$u^\alpha \nabla_\alpha u^\beta = 0. \tag{8.5}$$

The following lemma will have fundamental importance in General Relativity.

Lemma 8.1 *The Equations (8.4) and (8.5) imply that the tensor*

$$T^{\alpha\beta} = r u^\alpha u^\beta. \tag{8.6}$$

satisfies the conservation laws

$$\nabla_\alpha T^{\alpha\beta} = 0. \tag{8.7}$$

Proof It follows from $u^\alpha u_\alpha =$ constant that $u^\beta \nabla_\alpha u_\beta = 0$. The proof is then obtained by computing first $u_\beta \nabla_\alpha T^{\alpha\beta}$. □

Remark 8.2 *In a proper frame the components of the tensor (8.6) are*

$$T^{00} = r, \quad T^{0i} = T^{i0} = 0, \quad T^{ij} = 0. \tag{8.8}$$

8.2 *Perfect fluids*

The equations of motion for Newtonian fluids are generalized to relativistic fluids by introducing a spacetime energy momentum stress tensor T (the name abbreviated to stress energy tensor), such that its components T^{00}, T^{0i}, and T^{i0} in a proper frame of the fluid represent energy, energy flux, and momentum[24] relative to that frame, while the space part of this tensor is, in a first study, identified with the stress tensor of Newtonian mechanics.

A relativistic fluid is called a **perfect fluid** if the components of T in a *proper frame* reduce to:

$$T^{00} = \mu, \quad \text{energy density,} \tag{8.9}$$

$$T^{0i} = 0 \quad \text{no energy flux (no heat flow)} \tag{8.10}$$

$$T^{i0} = 0, \quad \text{zero momentum} \tag{8.11}$$

and, with p a scalar function called pressure and e the Euclidean metric

$$T^{ij} = p e^{ij}, \tag{8.12}$$

The formula (8.12) coincides with the stress tensor of a Newtonian perfect fluid[25]

[24] We do not give an axiomatic formulation for these physical notions: physics cannot be axiomatized (neither, by the way, can mathematics at its very foundations). Eventually energy and momentum are defined as mathematical objects in elaborate theories. For instance, energy appears as the conserved quantity associated with a time-independent hamiltonian.

[25] The force acting on an element of the boundary of an elementary volume is orthogonal to it: there is no antislipping force.

Proposition 8.3 *The stress energy tensor of a relativistic perfect fluid is in arbitrary coordinates*

$$T^{\alpha\beta} = \mu u^\alpha u^\beta + p(\eta^{\alpha\beta} + u^\alpha u^\beta) \tag{8.13}$$

Proof One checks that the above tensor has components in the proper frame given by Equations (8.9–12). □

Remark 8.4 *In the usual pressure and mass units, p is of the order of c^{-2} with respect to μ; hence at ordinary scales p is very small with respect to μ.*

The proper frame varies from point to point. It is not in general the natural frame of an inertial coordinate system. It is therefore not legitimate to write the usual equations of separate conservation of energy and momentum by using ordinary partial derivatives and the components (8.9–12) of the stress energy momentum tensor in a proper frame of the fluid. The dynamical equations of a perfect fluid are postulated to be spacetime tensorial equations which read in arbitrary coordinates

$$\nabla_\alpha T^{\alpha\beta} = 0. \tag{8.14}$$

These equations can be written in an arbitrary Lorentzian metric. We will return to them in the next chapter (General Relativity).

III

GENERAL RELATIVITY AND EINSTEIN'S EQUATIONS

1 Introduction

The gravitational field is the only universal force noticeable by everyone everywhere and known since antiquity. It prevents massive free bodies from moving in straight lines with constant velocities with respect to inertial frames. This result is known to a child throwing a ball, which describes approximately a parabola[1]. It was known to Tycho Brahé, who observed the trajectories of the planets around the Sun, and to Kepler who showed that these trajectories are ellipses. These observations were an inspiration for Newton to elaborate his theory of gravitation. Newton's gravitational law has proven to be in very good agreement with experiments on Earth and in the modelling of the motion of the solar planets, except for a small advance of the perihelion of Mercury unexplained by the perturbation due to other planets.

Einstein's genius discovered in 1915 the theory of General Relativity. It introduces a curved Lorentzian metric to represent spacetime to reconcile Special Relativity, whose invariance group is the Poincaré group, with the Newtonian gravity law. Indeed, Newton's spacetime is $E_3 \times R$ and the invariance group of Newton's gravity is the Galileo group. General Relativity coincides with Special Relativity when gravity can be neglected, and Einstein's theory of gravitation nearly coincides with the Newtonian theory in the physical circumstances where the latter has proved to be accurate; that is, when the velocities of the gravitating bodies are slow (compared with the speed of light) and the gravitational fields are relatively weak. Einstein's theory also predicted strange and new phenomena, such as the expansion of the Universe (Chapter 5), black holes (Chapters 13 and 14), gravitational lenses and gravitational waves, all of which have now been verified by direct or indirect observation.

Extensions to dimensions greater than 4 are now a subject of active study, in relation to the unification of the fundamental interactions and string and brane theories, with applications to astrophysics and cosmology.

2 Newton's gravity law

The fundamental law of Newtonian dynamics, $F = m\gamma$, relates the acceleration γ of a test particle with its "inertial" mass m and the force F acting on it. In its

[1] According to Newton it describes an ellipse with focus at the centre of the Earth.

primitive, Newtonian, form the equivalence principle is the expression of the fact that the acceleration due to gravity $\gamma(x,t)$ of a test[2] particle situated at a point (x,t) of spacetime is independent of the nature of this particle. The gravitational force in Newton's dynamics is

$$F_{grav} = m\gamma_{grav} \tag{2.1}$$

One used to interpret this formula by saying that the gravitational mass is equal to the inertial mass. This Newtonian equivalence principle was first stated by Galileo[3], then checked with greater and greater precision by Eötvos (1933), Dickie (1962), and later experiments.

Remark 2.1 *In the light of what we know of other fundamental fields (electromagnetic, Yang–Mills,...) one could say that m is the gravity charge, but if we take into account the results of Special Relativity we see that gravity is fundamentally different from other fields. The mass m appearing in dynamics is a constant attached to the particle only in a first approximation; it actually depends on its motion.*

In Newtonian theory the gravitational vector field γ_{grav} is the gradient of the Newtonian potential, a scalar function U on Newton's spacetime $E_3 \times R$. The fundamental law of Newtonian dynamics implies that the motion of a test particle in a gravitational field obeys a differential equation independent of its mass, which reads in inertial coordinates

$$\frac{d^2x^i}{dt^2} = \gamma^i_{grav} = \frac{\partial U}{\partial x^i}. \tag{2.2}$$

The gravitational potential U satisfies the Poisson equation:

$$\Delta U = -4\pi\kappa\rho \tag{2.3}$$

where Δ is the Laplace operator of Euclidean space E_3, ρ is a possibly time-dependent positive scalar function on R^3 equal to the mass density of the sources at time t, and κ is the Newtonian gravitational constant. The value of this universal constant depends on the choice of units for time, length, and mass. It is approximately, in CGS units, the very small number:

$$\kappa \cong 6.67259 \times 10^{-8} cm^3 g^{-1} s^{-2}. \tag{2.4}$$

We see from (2.3) that the Newtonian gravitational potential U is determined at each instant of time by the mass density at this time. Newton himself had some misgivings about this instantaneous "action at a distance", which surprises our human experience.

[2] That is, so small that it does not itself creates gravity.

[3] The story (but perhaps not history) is that he dropped various objects from the Leaning Tower of Pisa, and saw them reach the ground at the same time.

3 General relativity

Besides his desire to reconcile Special Relativity, valid when gravitational effects are negligible, and Newton's gravity law, Einstein was led to his theory by physical facts and new ideas.

3.1 *Physical motivations*

Two physical facts inspired Einstein.

- The **principle of general covariance**, an extension of the principle called "material indifference" in Newtonian mechanics, which is essentially to say that the physical phenomena themselves do not depend on the reference frame in which we express their laws. The laws may take different forms in different frames, but it should be possible to find for them frame-independent formulas. Tensor fields are good candidates as objects for physical laws, being intrinsic (geometric) objects on a manifold, represented by their components, specific numbers attached to them by the choice of a particular reference frame. We have already expressed the dynamics of Special Relativity in arbitrary coordinates. In the motion of test particles the inertial forces appear through the Christoffel symbols of the metric.
- The fact that the Newtonian gravitational acceleration is the gradient of a scalar field U on a given arena $E_3 \times R$. This geometrical fact expresses the Newtonian equivalence of inertial and gravitational mass.

In Einstein's General Relativity, gravity manifests itself by a tensor field and, in the absence of other forces, the motion of a test particle is determined by this tensor field; it does not depend on the mass of the particle. Einstein's revolutionary idea is that the arena where the gravity tensor lives is itself determined by gravity, both are united in a *4-dimensional Lorentzian manifold* (V, g), *called the* **spacetime**. The trajectories of test particles are geodesics of the metric g. Einstein's **weak equivalence principle** corresponds to the fact that in a general Lorentzian spacetime the geodesic equations are formally the same as in Minkowski spacetime with arbitrary coordinates where the non-vanishing of the Christoffel symbols signals the presence of inertial forces due to the non-inertial frame of reference. In a general spacetime (V, g), it is always possible at one given point (see the exercise in Chapter 1) to choose local coordinates such that the Christoffel symbols vanish at that point; gravity and relative acceleration are then, at that point, exactly balanced. It is even possible to choose local coordinates such that the Christoffel symbols vanish along one given geodesic; astronauts in spacecraft have made popular the fact that in free fall one feels neither acceleration nor gravity; in a small enough neighbourhood of a geodesic the relative accelerations of objects in free fall are approximately zero.

However, there is no intrinsic splitting between gravity and inertial type forces: in a general Lorentzian manifold there is no coordinate system (except if the metric is locally flat) in which all the Christoffel symbols vanish in the domain

of a chart, geodesics are not represented by straight lines. In the presence of a non-everywhere-vanishing mass or energy density, gravity manifests itself by the non-vanishing of the Riemann curvature tensor $Riemann(g)$. If $Riemann(g)$ does not vanish in an open set U of a spacetime (V, g), the gravitational field in U cannot be identified with an inertial field. The existence of a non-zero curvature of the metric is experienced, beyond the first approximation, by the relative acceleration of test particles, as is predicted by the equation of geodesic deviation (Section I.7).

To summarize, *in General Relativity space and time are united in a differentiable manifold V endowed with a pseudo-Riemannian metric g of Lorentzian signature. The length of a timelike curve measures the intrinsic, proper, time along that curve. Massive pointlike objects in free fall follow a timelike geodesic. Light rays follow null geodesics.*

4 Observations and experiments

4.1 *Deviation of light rays*

A consequence of Einstein's postulate of the equivalence of mass and energy is that point-like test particles follow geodesics of the spacetime metric. The light rays in particular are null geodesics: light is deflected by a gravitational field. This fact has been confirmed by observation as early as 1919 by Eddington and Dyson during a solar eclipse, though not with very high precision due to large experimental errors. Indeed, a solar eclipse makes possible a photograph of the sky where stars are visible, so that the position of their images can be compared with these positions in a photograph taken in the absence of the sun. Better precision is now obtained for the deflection of radio waves, which do not need an eclipse to be observed.

A spectacular verification of the deflection of light rays by a gravitational field is the observed phenomenon called a **gravitational lens**[4]. The light rays issued from a star or galaxy, i.e. the null geodesics issued from some point in the curved spacetime, may intersect again and an observer can see several images simultaneously. Several such gravitational lenses have now been observed. In particular the "Einstein cross", obtained by the Hubble telescope, consists of four images of the same quasar; in the centre appears a fainter image of a galaxy whose gravitational field acts as lens.

4.2 *Proper time, gravitational time delay*

As in Special Relativity the physically measurable quantity is **proper time**, which is the length in the spacetime metric of timelike curves. It is again postulated that the period of specific spectral rays of atoms provides universal clocks for the measure of proper time.

[4] See Schneider, P., Ehlers, J., and Falco, E. E. (1992) *Gravitational Lenses*. Springer-Verlag, Berlin.

A consequence of the reality of the dependence of this physically measurable time on the presence of a gravitational field is the observation of a **red shift** in the spectrum of a given atom in a gravitational field, as seen by a distant observer for which this field is weaker. Such shifts add to the Doppler effect for bodies in motion. This fact, the relative weakness of the gravitational field on Earth, and the imprecise knowledge of the gravitational field at the surface of stars did not permit in Einstein's time an accurate verification of the agreement of a theoretical estimate issued from Einstein's hypothesis, and observations. The **Mössbauer** resonant effect (emission of γ rays with an extremely narrow profile, reabsorbed in crystals with a very sharp resonance) enabled Pound and Rebka (1960) to measure the apparent shift of spectral lines due to the variation of the gravitational field with height over the Earth's surface in their laboratory. We give below a brief account of the theoretical prediction in this case.

Inspired by the Newtonian approximation (see Section 3.6) we suppose the spacetime metric to be, in the laboratory,

$$g = -N^2(x^1)dt^2 + g_{11}(x^1)(dx^1)^2 + g_{ab}dx^a dx^b, \quad a, b = 2, 3 \tag{4.1}$$

where x^1 varies with the height of a point while x^a, $a = 2, 3$ label the position of this point in a horizontal plane. Consider two identical clocks, one at rest on the ground, say $x^1 = 0$, the other at rest on the same vertical, with height labelled by $x^1 = h$. The hypotheses are therefore that, in the coordinates x^α, the representatives of the world lines of the clocks lie in the plane (x^1, t); they are the lines where t only varies, given respectively by $x^1 = 0$ and $x^1 = h$. Let $T(0)$ and $T(h)$ be the periods, in proper time, of these clocks. The signals of the beginning and end of a period of the clock on the ground are transmitted by light rays to an observer handling the other clock. Such light rays obey the differential equation

$$\frac{dx^1}{dt} = \frac{\sqrt{g_{11}}}{N}. \tag{4.2}$$

A signal emitted at parameter time $t(0)$ reaches the clock at height h at parameter time

$$t(h) = \int_0^h \frac{N}{\sqrt{g_{11}}}(x^1)dx^1 + t(0). \tag{4.3}$$

The difference in the parameter times between the two emitted and two received signals is therefore the same. However, the same is not true for the proper times, which are the physically measured quantities. We denote by $T(0)$ [respectively $T(h)$] the period in proper time of the clocks. Using the relation between the parameter time and the proper time for each of the clocks we find that:

$$t_2(0) - t_1(0) = N^{-1}(0)T(0) = t_2(h) - t_1(h) = N^{-1}(h)T(h) \tag{4.4}$$

The corresponding lapse of proper time marked by the clock of this observer is

$$T(h) = \frac{N(h)}{N(0)} T. \tag{4.5}$$

A longer period is observed (redshift) if $N(h) > N(0)$, which is the case in our laboratory example where the source of gravitation is the Earth with mass m; then (see Newtonian approximation in Section 3.6):

$$N^2 \cong 1 - 2\frac{m}{x^1}. \tag{4.6}$$

Theory and experimental results have proven to be in excellent agreement.

The dependence of proper time on the gravitational field has been measured directly for celestial objects and called **time delay** by Shapiro, who measured it first in 1966, by sending radar signals to planets (i.e. in the gravitational field of the Sun) and measuring the time elapsed on Earth between their emission and their return after reflection. The same experiment with more precise results has been done more recently with spacecraft orbiting and on the surface of Mars. We will give the computation of the theoretical prediction in Chapter 4.

The reality of the proper time defined by the Lorentz metric characterizing spacetime in the presence of a gravitational field has been checked directly by carrying caesium clocks on satellites around the Earth and observing that they have gained some hundred nanoseconds[5] over the same type of clock that stayed on the ground.

Finally, the Global Positioning System (GPS) needs for its accuracy the General Relativity correction of the physical (proper) time[6].

All experimental results to date[7] are in agreement with the predictions of Einstein's theory.

5 Einstein's equations

General Relativity replaces Newton's scalar potential U on E_3 by a symmetric 2 tensor g on a four-dimensional manifold V. The components $g_{\alpha\beta}$ of g in local coordinates are often called gravitational potentials. The pair (V, g) should be determined by the gravity sources.

The first problem, solved by Einstein, was to find an equation for the tensor g to replace Newton's equation for the scalar potential U.

5.1 *Vacuum case*

The vanishing in a domain of spacetime of the Riemann curvature tensor of the spacetime metric g implies that this metric is locally flat in this domain

[5] After correction of the kinematic effect of Special Relativity. A motorless satellite describes a geodesic of the Lorentzian (Schwarzschild) metric, the observer on Earth, not in free fall but held by the ground, does not.

[6] See for instance the article by Ashby in *Living Reviews*.

[7] For a description of experiments, their results and discussions, see for instance Ohanian, H. and Ruffini, R. (1994) *Gravitation and Spacetime*, 2nd edn, Ch. 4, Norton.

(cf. e.g. CB-DM1 V B 2), and hence without gravitational effects. It is therefore a condition too strong to impose in domains empty of energy sources (vacuum) in a non-globally empty spacetime. A natural candidate for equations to impose *in vacuo* on the Lorentzian metric g, a symmetric 2-tensor, is the vanishing of the symmetric 2-tensor $Ricci(g)$, which is related to the Riemann curvature tensor by contraction. The tensor $Ricci(g)$ is a second-order partial differential operator on g, like the Laplace equation for the Newtonian potential.

In an open set U devoid of energy sources the Lorentzian metric g of an Einsteinian spacetime[8] (V, g) satisfies the following tensorial **Einstein equations in vacuum**:

$$Ricci(g) = 0. \tag{5.1}$$

These equations are represented in each chart (see Equations I.7.6, I.6.13), with domain included in U, by the following system of second-order quasilinear[9] equations for the components $g_{\alpha\beta}$ of the metric

$$R_{\alpha\beta} \equiv \partial_\lambda \Gamma^\lambda_{\alpha\beta} - \partial_\alpha \Gamma^\lambda_{\beta\lambda} + \Gamma^\lambda_{\alpha\beta}\Gamma^\mu_{\lambda\mu} - \Gamma^\lambda_{\alpha\mu}\Gamma^\mu_{\beta\lambda} = 0. \tag{5.2}$$

These equations are not independent; they satisfy the identities deduced in Chapter I from the Bianchi identities:

$$\nabla_\alpha S^{\alpha\beta} \equiv 0, \qquad S^{\alpha\beta} \equiv R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R, \quad R \equiv g^{\lambda\mu} R_{\lambda\mu}. \tag{5.3}$$

The second problem, a subject of never-ending studies, is the search for a physically meaningful mathematical representation of sources of the gravitational field to replace the mass density ρ which appears on the right-hand side of the Poisson equation satisfied by the Newtonian potential.

5.2 *Equations with sources*

The choice of equations connecting the metric g with sources took a considerable amount of time for Einstein to settle on. Einstein was inspired on the one hand by the relativity to observers of the splitting between energy and momentum (see Equation II.7.2) and the conservation laws found in Special Relativity (see Equations II.3.12, II.8.14), and on the other hand by the Bianchi identities (I.8.6). After several years of unsuccessful attempts, Einstein, with the help of the mathematician Marcel Grossman, found the following **Einstein equations**, more than ever universally adopted today:

$$Einstein(g) + \Lambda g \equiv Ricci(g) - \frac{1}{2} g R(g) + \Lambda g = T. \tag{5.4}$$

[8] Classically the spacetime is of dimension 4, but higher dimensional spacetimes are also considered now in unified field theories.

[9] It is known now that the Einstein equations are the only tensorial ones which are second order and quasilinear for the metric g in four dimensions. There exist other possibilities (Gauss–Bonnet equations) in higher dimensions, which seem to play a role in string theory.

In local coordinates these take the form

$$S_{\alpha\beta} + \Lambda g_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R + \Lambda g_{\alpha\beta} = T_{\alpha\beta}, \tag{5.5}$$

where T is a symmetric 2-tensor, the stress energy tensor, supposed to represent the density of all the energies, momenta and stresses of the sources. It is a phenomenological tensor whose choice is not always easy, even in the classical dimension 4. Λ is a number, called the cosmological constant. It was not included in the original Einstein equations; he added it when looking for a stationary model for the cosmos, and then removed it[10] after the interpretation of astronomical observations as expansion of the universe. It is again included by cosmologists, but rather as a scalar field. It is generally considered to be very small now but to have been large in the early Universe.

The equations (5.4) on an $n + 1 = d$ dimensional spacetime are equivalent to the following ones

$$R_{\alpha\beta} = \rho_{\alpha\beta}, \quad \text{with} \quad \rho_{\alpha\beta} \equiv T_{\alpha\beta} + \left(\frac{d}{d-2} \Lambda - \frac{T^{\lambda}_{\lambda}}{d-2} \right) g_{\alpha\beta}. \tag{5.6}$$

The Einstein equations (5.4) together with the identities (5.3) imply that the stress energy tensor must satisfy the following equations, also called the conservation laws, or the equations of motion:

$$\nabla_{\alpha} T^{\alpha\beta} = 0. \tag{5.7}$$

We found in Chapter 2, for the simplest sources, a stress energy tensor which satisfies these equations in the Minkowski spacetime. Einstein postulated the following **strong equivalence principle**, which extends to all sources the identification of inertial and gravitational forces found for test particles:

Proposition 5.1 *The stress energy tensor for sources of the usual physical type in the presence of gravitation, that is in a curved Lorentzian spacetime, is formally the same as the stress energy tensor for this type of source in Special Relativity, with the spacetime metric g replacing the Minkowski metric. The equations of motion are obtained by using the covariant derivatives in the metric g. The Christoffel symbols replace both the gravitational and inertial forces of Newton's theory. They can be made to vanish at any spacetime point by choice of coordinates.*

We see that in Einstein's theory, in contrast with Newton's, the equations of motion are a consequence of the equation satisfied by the gravitational potential. Even the geodesic equation of motion of a test particle can be deduced from the conservation laws, by taking for the source a dust (see next section) with density given by a Dirac delta function supported by a time line.

[10] and called it "the greatest blunder of his life".

6 Field sources

There are, on a macroscopic scale[11], two type of source: field sources and matter sources. The properties of the former are readily deduced from their Special Relativistic expressions. The latter are more phenomenological and rest on a less sure footing already in Special Relativity. We will return to them in Chapters 9 and 10.

6.1 *Electromagnetic sources*

6.1.1 *Maxwell's equations*

The **electromagnetic field** on a spacetime (V, g) is the only one apparent at the macroscopic level. It is, as in Special Relativity, represented by a 2-form F on V. In the natural frame with $\wedge$ the exterior product of forms, one has

$$F \equiv \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta. \tag{6.1}$$

Maxwell's equations, deduced from their expression in Special Relativity and the equivalence principle are

$$dF = 0, \quad \text{and} \quad \nabla \cdot F = J, \tag{6.2}$$

with J a vector field on V, the $n+1$ electric current[12]. They read in coordinates like in Special Relativity and arbitrary coordinates

$$\nabla_\alpha F_{\beta\gamma} + \nabla_\gamma F_{\alpha\beta} + \nabla_\beta F_{\gamma\alpha} = 0, \quad \text{and} \quad \nabla_\alpha F^{\alpha\beta} = J^\beta. \tag{6.3}$$

They imply the conservation equation:

$$\nabla \cdot J = 0, \quad \text{i.e.} \quad \nabla_\beta J^\beta = 0. \tag{6.4}$$

Remark 6.1 *The second set of Maxwell's equations $\nabla \cdot F = J$ can also be written $\delta F = J$, using the L^2 adjoint δ (see Chapter 1) of the differential operator d.*

6.1.2 *Einstein–Maxwell equations*

The stress energy tensor of an electromagnetic field on a general spacetime has, according to the equivalence principle, the same expression as in Special Relativity,

$$\tau_{\alpha\beta} := F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}. \tag{6.5}$$

It is easy to see that in the case $n + 1 = 4$ the Maxwell tensor has a zero trace, $\tau^\alpha_\alpha \equiv 0$. This property is linked with the conformal invariance of Maxwell's equations in dimension 4 (see Appendix VI).

[11] At the elementary particle level, one calls matter the fermions. They are represented by spinor fields (see e.g. CB-DM1 Vbis, CB-DM2 IV 2), and quantum effects become relevant. At the cosmological scale "matter" seems more and more mysterious.

[12] Classically, there exist no magnetic charges; if there were the 2-form F would not be closed.

The following theorem holds, with the same proof as in Section II.3.

Theorem 6.2 *When Maxwell's equations are satisfied it holds that:*

$$\nabla_\alpha \tau^{\alpha\beta} = J^\lambda F^\beta{}_\lambda. \tag{6.6}$$

The vector $J^\lambda F^\beta{}_\lambda$ is called the Lorentz force, it vanishes in the absence of an electric current.

6.1.3 *Electrovac Einsteinian spacetimes*

Such spacetimes satisfy Maxwell's equations with $J = 0$ and the Einstein equations

$$S_{\alpha\beta} = \tau_{\alpha\beta}. \tag{6.7}$$

The Bianchi identities and the conservation law of the Maxwell tensor τ in the absence of an electric current make these equations compatible.

6.1.4 *Electromagnetic sources*

To a unit timelike vector field u, i.e. in physics language to an observer, is associated the electric vector field E, orthogonal to u:

$$E_\alpha = F_{\beta\alpha} u^\beta \tag{6.8}$$

Remark 6.3 *A magnetic vector field H is defined only when $n = 3$. In this case the Hodge adjoint of the 2-form F is a 2-form given by:*

$$({}^*F)_{\alpha\beta} = \frac{1}{2}\mu_{\alpha\beta\lambda\mu} F_{\lambda\mu}. \tag{6.9}$$

The magnetic vector field with respect to an observer with timelike unit vector u is then defined by

$$H_\alpha := ({}^*F)_{\beta\alpha} u^\beta. \tag{6.10}$$

The usual classical model for the electric current in a charged fluid with flow lines that are trajectories of a vector field u is:

$$J^\alpha \equiv q u^\alpha + \sigma E^\alpha \tag{6.11}$$

with q the charge density and σ the electric conductivity. The simplest cases of conductivity are $\sigma = 0$, realized in particular in vacuum, and $\sigma = \infty$, which implies, for J to be finite, the vanishing of the electric field E. This case is called **magnetohydrodynamics**.

Remark 6.4 *In the case $n = 3$ one may add to J a current due to the so-called* ***Hall effect*** *of the type:*

$$J_\alpha := \tau P_\alpha, \quad \text{with} \quad P_\alpha := \mu_{\alpha\beta\lambda\mu} u^\beta E^\lambda H^\mu \tag{6.12}$$

τ is the Hall coefficient and P is the Poynting vector. It is in the orthonormal frame with time axis u (proper frame of the fluid), the space vector given by

$$P = E \wedge H. \tag{6.13}$$

6.2 *Electromagnetic potential*

Since F is a closed 2-form there exists on any simply connected domain[13] a 1-form A defined up to the addition of an exact differential such that

$$F = dA, \quad \text{i.e.} \quad F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha. \tag{6.14}$$

Such a 1-form is called an electromagnetic potential.

Maxwell's equation $\delta F = J$ is a second-order system for the vector potential A. Since F can also be written[14] $F_{\alpha\beta} = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha$, we can write this Maxwell equation as the system

$$\nabla_\alpha(\nabla^\alpha A^\beta - \nabla^\beta A^\alpha) = J^\beta. \tag{6.15}$$

The equations (6.15) are not independent, since $\delta^2 \equiv 0$. One may take advantage of the indeterminacy of A to impose that it satisfies an additional equation, called a gauge condition. An early favoured one was the **Lorentz condition**

$$\delta A = 0, \quad \text{i.e.} \quad \nabla_\alpha A^\alpha = 0. \tag{6.16}$$

Under this condition the potential A satisfies a system of linear wave equations in the spacetime metric; namely, using the Ricci formula, one has

$$\nabla_\alpha \nabla^\alpha A^\beta - R^\beta{}_\lambda A^\lambda = J^\beta. \tag{6.17}$$

Remark 6.5 *These equations have the disadvantage of making second derivatives of the space time metric g appear. We will see in Chapter 6 that if we consider the components A^β of A as a set of scalar functions and if the local coordinates are harmonic, this set of Maxwell equations read as a system of wave equations for the functions A^β which contain only g and its first derivatives. An alternative, to have this result in arbitrary coordinates, is to make the gauge choice $\partial_\alpha A^\alpha = 0$ instead of $\nabla_\alpha A^\alpha = 0$.*

Other gauges are in use, in particular the **temporal gauge** $A_0 = 0$ and the **Coulomb gauge** $\nabla_i A^i = 0$.

6.3 *Yang–Mills fields*

The non-uniqueness of the electromagnetic potential A, the local character of its definition on a multiply connected manifold, and the physical properties of spinor fields in the presence of an electromagnetic field, have led to the interpretation of A as the representative on spacetime of a $U(1)$ connection on its tangent bundle. The electromagnetic field being the curvature of this connection.

After the discovery of other fundamental interactions, strong and weak, it was found by Yang and Mills that these interactions could also be mathematically modelled by curvatures of connections on the tangent bundle of spacetime, but with higher dimensional and non-Abelian groups. The electromagnetic and

[13] Poincaré lemma, see for instance CB-DM1, IVbis B.
[14] Due to the symmetry of the Christoffel symbols.

weak interactions were unified by Weinberg and Salam as the curvature of an $SU(2)\times U(1)$ connection, with the addition of a scalar field, the Higgs field, made necessary by the physical fact of the distinction (called symmetry breaking) of these interactions at lower energy scale. We recall briefly[15] the descriptions of a classical (non-quantized) Yang–Mills field, the Yang–Mills equations and a conserved stress energy tensor which can be taken as a source[16] for the Einstein equations.

Let G be a Lie group and $\mathcal{G}$ be its Lie algebra. A representative on spacetime of a G connection is a locally defined 1-form with values in $\mathcal{G}$ called a Yang–Mills potential. The curvature of this connection is represented by a $\mathcal{G}$ valued 2-form, called a Yang–Mills field, given in terms of A by (with [.] the Lie bracket in $\mathcal{G}$)

$$F = dA + [A, A], \quad \text{i.e.} \quad F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta]. \tag{6.18}$$

The 2-form F satisfies the identity

$$\hat{d}F \equiv 0 \quad \text{i.e.} \quad \hat{\nabla}_\alpha F_{\beta\gamma} + \hat{\nabla}_\gamma F_{\alpha\beta} + \hat{\nabla}_\beta F_{\gamma\alpha} \equiv 0, \tag{6.19}$$

where $\hat{d}$ is the gauge covariant exterior differential and $\hat{\nabla}$ is the metric and gauge covariant derivative:

$$\hat{\nabla}_\alpha F_{\beta\gamma} := \nabla_\alpha F_{\beta\gamma} + [A_\alpha, F_{\beta\gamma}]. \tag{6.20}$$

The Yang–Mills equations consist of the identity (6.19) and the equations generalizing the second set of Maxwell equations

$$\hat{\nabla}_\alpha F^{\alpha\beta} \equiv \nabla_\alpha F^{\alpha\beta} + [A_\alpha, F^{\alpha\beta}] = 0. \tag{6.21}$$

Remark that (except for Abelian groups G) the Yang–Mills potential does not disappear from these equations.

The Yang–Mills stress energy tensor, a generalization of the Maxwell tensor, is given by

$$\tau_{\alpha\beta} := F_\alpha{}^\lambda \cdot F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} \cdot F_{\lambda\mu}. \tag{6.22}$$

where a dot denotes a scalar product with respect to the Killing form[17] of $\mathcal{G}$. It is divergence free if the Yang–Mills equations are satisfied.

Remark 6.6 *On a d-dimensional spacetime the Einstein equations with Yang–Mills source are equivalent to*

$$R_{\alpha\beta} = \rho_{\alpha\beta} \equiv F_\alpha{}^\lambda \cdot F_{\beta\lambda} - \frac{1}{2(d-2)} g_{\alpha\beta} F^{\lambda\mu} \cdot F_{\lambda\mu}. \tag{6.23}$$

The Yang–Mills stress energy tensor has a zero trace, as the Maxwell tensor, in four-dimensional spacetimes; in this case $\rho_{\alpha\beta} = \tau_{\alpha\beta}$.

[15] See for instance CB-DM1 Vbis Problem 1.
[16] Though, at macroscopic scale, the Yang–Mills fields are not directly observed.
[17] See for instance CB-DM1 III D 6.

6.4 *Scalar fields*

6.4.1 *Real scalar fields*

In modern cosmology one introduces on the spacetime (V, g) a real scalar field ψ with potential U, a function of ψ. The stress energy tensor of a real scalar field is

$$T_{\alpha\beta} \equiv \partial_\alpha\psi\partial_\beta\psi - \frac{1}{2}g_{\alpha\beta}\partial^\lambda\psi\partial_\lambda\psi - g_{\alpha\beta}U(\psi). \tag{6.24}$$

It results from elementary computation that:

$$\nabla_\alpha T^{\alpha\beta} \equiv \{\nabla^\alpha\partial_\alpha\psi - U'(\psi)\}\partial^\beta\psi, \quad U'(\psi) := \frac{dU}{d\psi}. \tag{6.25}$$

Therefore the divergence $\nabla\cdot T$ vanishes if and only if ψ is solution of the semilinear wave equation in the spacetime metric g,

$$\nabla^\alpha\partial_\alpha\psi - U'(\psi) = 0. \tag{6.26}$$

A cosmological constant Λ can be considered as a particular scalar field, with potential equal to Λ.

Remark 6.7 *The trace of T on a $d := n + 1$ dimensional spacetime is*

$$T^\lambda_\lambda \equiv \partial^\lambda\psi\partial_\lambda\psi\left(\frac{2-d}{2}\right) - dU(\psi) \tag{6.27}$$

therefore (see 5.6) the identity (6.24) is equivalent to

$$R_{\alpha\beta} = \partial_\alpha\psi\partial_\beta\psi + \frac{d-3}{2}g_{\alpha\beta}\partial^\lambda\psi\partial_\lambda\psi + \frac{d}{d-2}g_{\alpha\beta}U(\psi). \tag{6.28}$$

6.4.2 *Complex scalar field*

In unified models of fundamental interactions one introduces, besides a Yang–Mills field with potential A, a complex gauge covariant scalar field Ψ, with stress energy tensor

$$T_{\alpha\beta} = \frac{1}{2}\{D_\alpha\Psi(D_\beta\Psi)^* + (D_\alpha\Psi)^*D_\beta\Psi - g_{\alpha\beta}D^\lambda\Psi(D_\lambda\Psi)^*\}, \tag{6.29}$$

where a star denotes complex conjugation and D denotes the gauge covariant derivative

$$D_\alpha\Psi = \partial_\alpha\Psi + iA_\alpha\Psi. \tag{6.30}$$

6.5 *Wave maps*

In the reduction of d-dimensional Einstein equations under isometry groups one discovers certain fields which are mappings from the spacetime into some given Riemannian or pseudo-Riemannian manifolds and satisfy a so-called wave map equation. These mappings generalize the harmonic maps between Riemannian manifolds. A wave map is obtained as follows.

Let u be a smooth map between two pseudo Riemannian manifolds (V, g) and (M, h). Let x^α, $\alpha = 0, 1, \ldots, n$ be coordinates in an open set $\omega \subset V$, small enough for $u(\omega)$ to take its value in a domain of coordinates y^A, $A = 1, \ldots, d$ of the target M. The mapping u is represented in ω by d functions u^A of the $n+1$ variables x^α

$$x^\alpha \mapsto y^A = u^A(x^\alpha).$$

The gradient $\partial u(x)$ of u at $x \in V$, represented by $(\partial_\alpha u^A)$, is an element of the tensor product of the cotangent space to V at x by the tangent space to M at $u(x)$:

$$\partial u(x) \in T_x^* V \otimes T_{u(x)} M.$$

The gradient itself, ∂u, is a section of the vector bundle E with base V and fibre $E_x \equiv T_x^* V \otimes T_{u(x)} M$ at x. We endow E with a connection whose coefficients acting in $T_x^* V$ are the coefficients of the Riemannian connection at x of the metric g while the coefficients acting in $T_{u(x)} M$ are the pull back by u of the Riemannian connection 1-form of the metric h at $u(x)$. The **covariant differential** $\hat{\nabla} f$ of a section f of the fibre bundle E with base V is thus a section of the vector bundle $T_* V \otimes E$ over V. In local coordinates on V and M where the section f, for instance ∂u, is represented by the $(n+1) \times d$ functions f_α^A of the $n+1$ coordinates x^α, the covariant differential $\hat{\nabla}$ is represented by the $(n+1)^2 \times d$ functions

$$\hat{\nabla}_\alpha f_\beta^A(x) \equiv \partial_\alpha f_\beta^A(x) - \Gamma_{\alpha\beta}^\mu(x) f_\mu^A(x) + \partial_\alpha u^B(x) \Gamma_{BC}^A(u(x)) f_\beta^C(x).$$

where $\Gamma_{\alpha\beta}^\mu$ and Γ_{BC}^A denote respectively the components of the Riemannian connections of g and h.

Remark. The covariant differential $\hat{\nabla} g \equiv \nabla g$ of the metric g, section of $\otimes^2 T^* V$, is zero by the definition of its Riemannian connection. The field $h(u)$ defined by the mapping u and the metric h is a section of the vector bundle over V with fibre $\otimes^2 T_{u(x)}$ at x; it has also a zero covariant derivative $\hat{\nabla} h$, pull back by u of the Riemannian covariant derivative of h on M.

The Leibniz rule for the derivation of tensor products leads to the definition of **higher order covariant derivatives** sections of bundles over V with fibre $\otimes^p T_x^* V \otimes^q T_{u(x)} M$.

Ricci identity. Commutation of covariant derivatives gives, as could be foreseen, the following useful generalization of the Ricci identity:

$$(\hat{\nabla}_\alpha \hat{\nabla}_\beta - \hat{\nabla}_\beta \hat{\nabla}_\alpha) f_\lambda^A = R_{\alpha\beta\lambda}{}^\mu f_\mu^A + \partial_\alpha u^C \partial_\beta u^B R_{CB}{}^A{}_D f_\mu^D.$$

where $R_{\alpha\beta,\lambda\mu}$ and R_{ABCD} denote respectively the Riemann tensors of g and h.

A mapping u is said to be a **harmonic map** (a **wave map** if g is Lorentzian) if the g trace of the covariant derivative of its gradient vanishes. The equation

satisfied by such a map reads in local coordinates:

$$g^{\alpha\beta}\hat{\nabla}_\alpha\partial_\beta u^A \equiv g^{\alpha\beta}\{\partial_\alpha\partial_\beta u^A - \Gamma^\mu_{\alpha\beta}(x)\partial_\mu u^A + \Gamma^A_{BC}\partial_\alpha u^B\partial_\beta u^C\} = 0.$$

The **stress energy tensor** T of a wave map is

$$T_{\alpha\beta} = \partial_\alpha u.\partial_\beta u - \frac{1}{2}g_{\alpha\beta}\partial^\lambda u \cdot \partial_\lambda u \tag{6.31}$$

where the dot denotes the scalar product[18] in the metric h. The tensor T is divergence free.

Remark 6.8 *The energy of a wave map deduced from its stress energy tensor is not to be mistaken with what is called the energy of a harmonic map with source a properly Riemannian manifold (V, g), also called the Dirichlet energy, given by*

$$\int_V g(\partial u, \partial u)\mu_g. \tag{6.32}$$

6.6 *Energy conditions*

For the convenience of the reader we give here definitions usual in the literature of General Relativity.

Sources are said to satisfy:

1. The **weak energy condition** if the stress energy tensor is such that
$$T_{\alpha\beta}X^\alpha X^\beta \geq 0 \quad \text{for all timelike vectors } X. \tag{6.33}$$
2. The **strong energy condition** (also called Ricci positivity condition) if
$$\rho_{\alpha\beta}X^\alpha X^\beta \geq 0 \quad \text{for all timelike vectors } X, \quad \rho_{\alpha\beta} := T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T^\lambda_\lambda. \tag{6.34}$$
3. The **dominant energy condition** if the stress energy tensor is such that the vector $-T^\beta_\alpha X^\alpha$ is timelike and future directed for all timelike future directed vectors X.

7 Lagrangians

To save space and time, we use in this section a mathematically not very elaborate language.

We denote by μ_g the volume element of the metric g (see Section I.5.4).

7.1 *Einstein–Hilbert Lagrangian*

It seems that the following theorem was found in part independently by Einstein and Hilbert[19].

[18] Positive quadratic form if h is properly Riemannian.

[19] See the historical discussion in Damour, T. (2005) "Si Einstein m'était conté" Cherche midi 1905, translated by E. Novak as "Once upon Einstein", A. K. Peters, Ltd, Wellesley MA.

Theorem 7.1 *A metric g which is a solution of the Einstein equations in vacuo is a critical point, for variations δg with compact support, of the Einstein–Hilbert Lagrangian, defined by the scalar curvature $R(g)$*

$$L_{grav}(g) := \int R(g)\mu_g. \tag{7.1}$$

Proof We use as in Section 1.9 physicists' notation for tangent linear maps: δL, first variation of L, is the action of the Frechet functional derivative of L on the vector δg. It holds that

$$\delta L_{grav} = \int \delta R(g)\mu_g + \int R(g)\delta\mu_g. \tag{7.2}$$

The equations in Section I.9 show that if δg has compact support, then it holds that:

$$\int \delta R(g)\mu_g = \int R_{\alpha\beta}\delta g^{\alpha\beta}\mu_g \tag{7.3}$$

The classical formula for the derivative of a determinant gives

$$\delta\mu_g = \frac{1}{2}g^{\alpha\beta}\delta g_{\alpha\beta}\mu_g = -\frac{1}{2}g_{\alpha\beta}\delta g^{\alpha\beta}\mu_g. \tag{7.4}$$

Therefore we find that

$$\delta L_{grav}(g) \equiv \int \left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\right)\delta g^{\alpha\beta}\mu_g = -\int S^{\alpha\beta}\delta g_{\alpha\beta}\mu_g. \tag{7.5}$$

The (Frechet) derivative of L_{grav} vanishes at g if and only if δL_{grav} vanishes for any set of $\delta g^{\alpha\beta}$, i.e. if and only if the vacuum Einstein equations are satisfied. □

7.2 *Lagrangians and stress energy tensors of sources*

The Lagrangians of sources are deduced from their expressions in Special Relativity and the equivalence principle. It is easy to check that the field equations of Section 6 are the Euler equations of Lagrangians (i.e. that the solutions of these field equations are critical points of these Lagrangians).

Theorem 7.2

1. *The Maxwell equations in vacuo are the Euler equations with respect to a closed 2-form F of the Lagrangian*

$$L_{e.m}(F, g) := \int F^{\alpha\beta}F_{\alpha\beta}\mu_g. \tag{7.6}$$

2. *The Lagrangian of a massless scalar field φ is*

$$L_{scal}(\varphi, g) := \int g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi\mu_g. \tag{7.7}$$

3. *The Lagrangians for Yang–Mills fields and wave maps are obtained by formulae respectively analogous to (7.6) and (7.7) by replacing ordinary products by scalar products: in the Lie algebra in the case of Yang–Mills, and in the metric of the target in the wave map case.*

Proof Straightforward calculations. □

7.3 *Coupled Lagrangian*

The following theorem legitimates the consideration of the sum of the Einstein–Hilbert Lagrangian and a source field Lagrangian as a Lagrangian for the Einstein equations with such sources.

Theorem 7.3 *Denote, generically, the field sources by ψ, and suppose that the field equations on spacetime derive from a Lagrangian invariant by diffeomorphisms, of the form*

$$L_{sour}(g,\psi) := \int l(g,\psi)\mu_g. \tag{7.8}$$

with first variation

$$\delta L_{sour} \equiv \int \{\Psi(g,\psi)\cdot\delta\psi + T^{\alpha\beta}(g,\psi)\delta g_{\alpha\beta}\}\mu_g, \tag{7.9}$$

where $\Psi = 0$ are the field equations, the dot denoting some algebraic linear form. Then

1. *The tensor T, with components $T^{\alpha\beta}$, is the stress energy tensor of the field ψ.*
2. *T satisfies the conservation law $\nabla\cdot T = 0$.*
3. *The vanishing of the variation with respect to g of the sum of the Einstein–Hilbert Lagrangian and the sources Lagrangian gives the Einstein equations with source the tensor T.*

Proof

1. Easy calculations show that the functional partial derivatives with respect to g of the Lagrangians (7.6) and (7.7) are respectively equal to the Maxwell stress energy tensor and to the stress energy tensor of a scalar field φ.
2. If the variation of g comes from a 1-parameter group of diffeomorphisms generated by a vector field X then

$$\delta g^{\lambda\mu} = (\mathcal{L}_X g)^{\lambda\mu} \equiv \nabla^\lambda X^\mu + \nabla^\mu X^\lambda. \tag{7.10}$$

If $\delta L = 0$ under such variations of g, and if $\Psi = 0$, then integration by parts shows that T has a vanishing divergence.

3. When the field ψ satisfies the field equations $\Psi = 0$, the first variation of the Lagrangian $L_{grav} + L_{sour}$ reduces to

$$\delta(L_{grav} + L_{sour}) = \int (-S^{\alpha\beta} + T^{\alpha\beta})\delta g_{\alpha\beta}\mu_g. \tag{7.11}$$

□

8 Fluid sources

As in Special Relativity a fluid source in a domain of a spacetime (V, g) is such that there exists in this domain a unit timelike vector field u, satisfying $g(u, u) \equiv g_{\alpha\beta}u^\alpha u^\beta = -1$, whose trajectories are the flow lines of matter. A moving Lorentzian orthonormal frame is called a proper frame if its timelike vector is u.

According to the strong equivalence principle the stress energy tensor of a **perfect fluid**, directly taken from its expression in Special Relativity, is, with μ and p respectively the energy and pressure densities

$$T^{\alpha\beta} := \mu u^\alpha u^\beta + p(g^{\alpha\beta} + u^\alpha u^\beta). \tag{8.1}$$

The equations satisfied by T are:

$$\nabla_\alpha T^{\alpha\beta} \equiv u^\beta \nabla_\alpha[(\mu + p)u^\alpha] + (\mu + p)u^\alpha \nabla_\alpha u^\beta + \partial^\beta p = 0. \tag{8.2}$$

They give, by contraction with the unit vector u, the continuity equation

$$\nabla_\alpha[(\mu + p)u^\alpha] - u^\beta \partial_\beta p = 0 \tag{8.3}$$

and the equations of motion, whose contraction with u vanishes,

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\beta u^\alpha)\partial_\alpha p = 0. \tag{8.4}$$

These equations are in general completed by an **equation of state** which reads for **barotropic fluids**

$$p = f(\mu). \tag{8.5}$$

Remark 8.1 *Equation (8.4) generalizes Newton's equation of motion. Equation (8.3) is not a conservation equation due to the term $u^\beta \partial_\beta p$, representing the work of pressure forces. This non-conservation was expected from the equivalence of mass and energy.*

Dust, also called incoherent matter, is the case of zero pressure.

More elaborate treatments, including thermodynamic considerations, will be given in Chapter 9. We will also consider the case of kinetic sources in Chapter 10.

A Lagrangian which depends on the choice of a special frame (flow lines as timelines), has been given by Taub for perfect fluids.

It seems that when dissipative effects are included, no Lagrangian in the usual sense exists.

9 Einsteinian spacetimes

9.1 *Definition*

A spacetime of General Relativity is a pair (V, g), with V a differentiable manifold and g a Lorentzian metric on V, both *a priori* arbitrary[20]. Such a spacetime is called Einsteinian if there exists a physically[21] meaningful stress energy 2-tensor T such that the following equations are satisfied on V, in whatever precise mathematical[22] meaning we can give to them:

$$Einstein(g) = T, \quad \text{i.e.} \quad S_{\alpha\beta} = T_{\alpha\beta} \tag{9.1}$$

and

$$\nabla \cdot T = 0, \quad \text{i.e.} \quad \nabla_\alpha T^{\alpha\beta} = 0. \tag{9.2}$$

The spacetime is called a vacuum spacetime if $T \equiv 0$.

Two isometric[23] spacetimes (V, g) and (V', g') are considered as physically identical.

Even in the vacuum case there exist many Einsteinian spacetimes. They have underlying manifolds with different topologies, as well as different metrics on each of these manifolds. In the following chapters, we will study the local and global constructions of Einsteinian spacetimes, in vacuum or with various sources. Some of these spacetimes have, at least at present, a purely mathematical interest, but a few of them are models of physical situations, at different time or space scales[24]. There is no universal Einsteinian spacetime as a model for reality; this is in complete disagreement with Newton's concepts, and also with Special Relativity.

Einstein's theory has, up to now, passed victoriously all experimental or observational tests.

9.2 *Regularity hypotheses*

It is known that, given a C^k, $k \geq 1$, structure on a manifold there always exists a C^∞ structure on V which induces this C^k structure. One does not restrict the generality of physical spacetimes by supposing that their manifold V is C^∞. In practice the **manifold** V support of the spacetime (V, g) is always supposed to be smooth, that is of class C^k with k large enough so that all the statements which are made concerning tensorial quantities, in the considered context, can be expressed in terms of their representatives in local coordinates. In practice, especially in the case of Einstein equations with isolated sources, one may find solutions in adjacent domains in different coordinate systems, and one wants to construct a solution in the union of these domains. One may be confronted then

[20] This is the origin of the name "General Relativity".
[21] This definition is, of course, ambiguous.
[22] See next section.
[23] That is, such that there exists a diffeomorphism f of V onto V' such that $f * g' = g$.
[24] The hypothesis that Einstein's equations are valid at different scales, from microscopic to cosmological, is however somewhat puzzling.

with the so-called problem of "matching conditions" to obtain such a solution. It is clear that geometrically (or physically) defined coordinates will have the best chance of leading to such a global solution. We will see a first example in the case of the Schwarzschild exterior and interior metrics.

If a change of coordinates is singular we consider that the new coordinates define another manifold. It is an extension of the region of the original manifold where the change of coordinates is smooth. We will see illustrations of this definition in Chapter 4.

What are the admissible function spaces on V where g must lie in order that the Einstein equations have a meaning? With the smoothness hypothesis made on V, hypotheses on the regularity of g are equivalent, in the domain of a chart, to hypotheses on the regularity of its components. A spacetime is of class C^k if the representations of the metric g in charts as well as its contravariant associate $g^{\$}$ are of class C^k. If $k \geq 2$, the tensor $Ricci(g)$ is classically defined, and is of class C^{k-2}. The class C^k, $k \geq 2$, spacetimes are a too restricted class to cover all interesting situations.

The definition of the Einstein equations, i.e. of the Ricci tensor, in the usual framework of Sobolev–Schwartz distribution theory is not straightforward, due to its non-linearity. A broad class of spacetimes defined in this well-known framework are those with **local $H_s(V)$ metrics**. The Sobolev spaces[25] W_s^p have proved a basic tool in the study of non-linear differential systems on R^n. The H_s spaces are particularly useful for mathematical (they are Hilbert spaces) and physical (they appear naturally in computations of physical energies) reasons. One says that a tensor T on V is in $H_s^{loc}(V)$ if T is in H_s on any compact subset of R^n contained in the image of a chart of V.

Lemma 9.1 *If the metric g is in $H_s^{loc}(V)$ as well as its contravariant associate $g^{\$}$, with $s > \frac{n+1}{2}$ and $s \geq 2$, then the Christoffel symbols and Riemann and Ricci tensors are almost everywhere defined tensors*[26] *such that*

$$\Gamma(g) \in H_{s-1}^{loc}(V), \quad Rieman(g), \quad Ricci(g) \in H_{s-2}^{loc}(V). \tag{9.3}$$

Proof The embedding and multiplication properties of Sobolev spaces[27], and the fact that the Riemann tensor is linear in second derivatives of g, quadratic in first derivatives of g, with coefficients polynomials in g and $g^{\$}$. □

In the proof of the existence and uniqueness of generic spacetimes from initial data (see Chapter 6) one splits space and time. Then $V = M \times R$, with M an n-dimensional manifold, and one introduces function spaces of tensor fields which admit restrictions to each $M_t := M \times \{t\}$ that are in a space $H_s^{loc}(M)$. The hypothesis $s > \frac{n}{2}$, $s \geq 2$, hence $s = 2$ if $n = 3$, is then sufficient for the

[25] See Appendix I.

[26] More general cases, with a non-integer s can be considered, but they are outside the scope of this book.

[27] See Appendix I and, for instance CB-DM2.

restriction to M_t of the Riemann tensor to be an almost everywhere defined function. Solutions of the Einstein equations with Ricci tensors equal to Dirac measures on submanifolds (and thin shells sources[28]) have also been considered.

To have global formulations on M it is convenient to endow it with a smooth Riemannian metric e and use the properties of the globally defined spaces $H_s(M, e)$ which are Hilbert spaces (see Appendix A).

We shall return later to these questions.

Other spaces are used in the study of shocks (see Chapter 9).

More refined spaces are presently used (Besov spaces, spaces with finite total variation,...), but their use is outside the scope of this book.

10 Newtonian approximation

We said that the Einsteinian theory of gravitation must nearly coincide with the Newtonian one in the physical circumstances where the latter has proved to be accurate. These circumstances are slow (compared with the speed of light) velocity of the gravitating bodies with respect to Newton spacetime and weak gravitational fields. In Einstein's theory the spacetime coincides with Minkowski spacetime in the absence of gravitation, that is with the manifold $R^3 \times R$ endowed with the flat metric, which reads in inertial coordinates and geometrical units as follows

$$-(dx^0)^2 + \sum_{i=1}^{3} (dx^i)^2.$$

An Einsteinian spacetime with a weak gravitational field will be $R^3 \times R$ with a metric such that

$$g_{00} = -(1 + f_{00}), \quad g_{0i} = f_{0i}, \quad g_{ij} = \delta_{ij} + f_{ij}$$

with $f_{\alpha\alpha}$ of order $\epsilon << 1$ while $f_{\alpha\beta}$ is of order ϵ^2 if $\alpha \neq \beta$.

In a spacetime with masses moving slowly (compared with 1, the speed of light) the derivatives $\partial_0 f_{\alpha\beta}$ will be small with respect to the space derivatives $\partial_i f_{\alpha\beta}$; these are considered as of the same order as $f_{\alpha\beta}$.

10.1 *Equations for potentials*

The component R_{00} in the considered approximation is equivalent to

$$R_{00} \cong \partial_i \Gamma^i_{00} \cong \frac{1}{2} \Delta f_{00},$$

with Δ the Laplace operator of Euclidean space. In Newtonian theory the gravitational field is the gradient of the Newtonian potential, a function U which satisfies the Poisson equation:

$$\Delta U = -4\pi\kappa\rho \tag{10.1}$$

[28] First by Taub, A. H. (1957) *Illinois J. Math.*, **1**, 370–88, and then Papapetrou, A. and Hamoui, A. (1967) *Ann. Inst. Poincaré A*, **VI**, 343–64.

where ρ is a positive scalar function on R^3 equal to the mass density of the sources and κ is the Newtonian gravitational constant. The value of this universal constant depends on the choice of units for time, length, and mass. As we said before, κ is approximately in CGS units 6.67259×10^{-8} cm^3 g^{-1} s^{-2}.

In Einstein's theory the energy sources are represented by the stress energy tensor. Under normal circumstances the most important energy source is pure matter, i.e. of the form with μ some positive scalar function:

$$T_{\alpha\beta} = \mu u_\alpha u_\beta.$$

In the case where the matter has a small velocity (with respect to the speed of light) for observers following the time lines of the coordinate system, it holds that u_0 is equivalent to -1 and u_i to zero; therefore T_{ij} is negligible and

$$\rho_{00} \equiv T_{00} - \frac{1}{2} g_{00} T \cong \frac{1}{2}\mu. \tag{10.2}$$

We see that in the approximation we are making, called the **Newtonian approximation**, the equation $R_{00} = \rho_{00}$ gives

$$\Delta f_{00} \cong \mu;$$

that is the Poisson equation of Newton's theory for the gravitational potential if f_{00} is identified with $-2U$ and μ with $8\kappa\pi$.

For this reason many physicists write the Einstein equations

$$S_{\alpha\beta} = 8\pi\kappa T_{\alpha\beta} \tag{10.3}$$

with $T_{\alpha\beta}$ the general relativistic phenomenological stress energy tensor.

The formula (10.3) mixes Einsteinian and Newtonian mechanics, since no unit of mass can be defined in Einstein's theory by choosing a standard kilogram. In fact a geometric unit of mass can be defined through the use of the universal gravitation constant: it is the equations

$$S_{\alpha\beta} = T_{\alpha\beta}, \quad T_{\alpha\beta} \equiv \mu u_\alpha u_\beta \tag{10.4}$$

which define the mass density of the dust creating the Einstein tensor $S_{\alpha\beta}$. We want to link this density with the Newtonian approximation density. We use the value of κ to find the geometric density μ corresponding to a Newtonian density ρ given in kilograms per cubic metre; it is given in inverse square of seconds by

$$\mu_{s^{-2}} \cong \pi 6.67259 \times 10^{-11} \rho_{\mathrm{kg/m^3}}. \tag{10.5}$$

We see that in Einstein's theory the mass density has as dimension the inverse square of time.

The mass of a body is the three-dimensional space integral of its mass density. Therefore the geometric mass m is given in cubic metres per second in terms of the Newtonian mass M, given in kilograms by

$$m_{\mathrm{m^3/s^{-2}}} \cong 8\pi 6.67259 \times 10^{-11} M_{\mathrm{kg}}. \tag{10.6}$$

Recall that in Relativity the speed of light is a universal constant, equal to 1 by definition in geometrical units. Therefore the mass of a body has in Einstein's theory the dimension of a length[29], equivalently the dimension of time.

To obtain in kilometres a mass given in kilograms one uses the Newtonian value of the speed of light, about $2.99792458 \times 10^8\,\mathrm{m\,s^{-1}}$. Hence in geometric units one has

$$m_{\text{meters}} \cong 8\pi 6.67259 \times 10^{-27} \times \{2.99792458\}^{-1} m_{\text{kg}}. \tag{10.7}$$

Example 10.1 *The formula gives for the Earth, with mass* $M_{\text{earth, kg}} = 5.98 \times 10^{24}{}_{\text{kg}}$:

$$m_{\text{earth, meters}} \cong 8\pi 6.67259 \times \{2.99792458\}^{-2} \times 5.98 \times 10^{-3} m_{\text{kg}}. \tag{10.8}$$

This is a geometric mass of about 1 centimetre.

In fact, the international unit of mass is not officially defined through the gravitational constant κ, because the value of this constant deduced from experiments (made in a Newtonian framework, as light was measured with pre-relativistic units) is not known with a great accuracy, due to the extreme weakness of the gravitational force between bodies of laboratory size. Another difficulty appears with masses of planetary size, which is that only the product of κ by their mass appears in the gravity law, and it is eventually their mass, not known otherwise, which is determined through this law when κ is known. Up to now in the International System of units the standard kilogram, deposited in Sèvres, still defines the unit of mass. It will probably soon be replaced by a unit defined through the Planck constant, which is measured to a very great accuracy, and the relation of energy (i.e. mass) with frequency issued from quantum machanics.

In this book, however, we write the Einstein equations in geometric units, i.e. $S_{\alpha\beta} = T_{\alpha\beta}$, introducing the relevant scale factor only when necessary for interpretation of observations.

10.2 *Equations of motion*

The geodesic equation governing the motion of a test particle in General Relativity is

$$\frac{d^2x^\alpha}{ds^2} + \Gamma^\alpha_{\mu\lambda}\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds} = 0.$$

For small velocities we have already seen that $\dfrac{dx^0}{ds} \sim 1$, $\dfrac{dx^i}{ds} \sim 0$; therefore

$$\frac{d^2x^i}{(dx^0)^2} \sim -\Gamma^i_{00} \sim -\frac{1}{2}\partial_i f_{00}. \tag{10.9}$$

[29] This result is in agreement with consideration of the metric as dimensionless and of the partial derivatives $\partial/\partial x^\alpha$ as with dimension L^{-1}, hence of $S_{\alpha\beta}$ of dimension L^{-2}. This interpretation is linked with viewing the coordinates, which in the fully Einsteinian theory are just parameters, as having the dimension of a length.

Einstein's law coincides in the Newtonian approximation, where $f_{00} = -2U$, with Newton's law.

11 Gravitational waves

An important feature of Einsteinian gravitation is its prediction of the propagation of the gravitational field with the speed of light (see Chapter 6). One speaks of "gravitational waves" and "gravitational radiation", though this last term has only relative meaning (for instance asymptotically in an asymptotically Minkowskian spacetime), since there exists no localized gravitational energy. On the other hand perturbations of a gravitational field, represented by a given Lorentzian manifold (V, g), can be intrinsically defined as tensor fields on the manifold V.

Gravitational waves are phenomena of very active theoretical and experimental research (see T. Damour "General Relativity Today"[30]).

11.1 *Minkowskian approximation*

We consider here weak gravitational fields, perturbations of the Minkowski spactime M^{n+1}. Specifically, taking inertial coordinates, we consider the metric with components on R^{n+1}

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{11.1}$$

with $\eta_{\alpha\beta} = diag(-1, 1, \ldots, 1)$ and suppose $h_{\alpha\beta}$ and its derivatives to be small[31] with respect to 1. It results from the expressions of the Ricci tensor and the Christoffel symbols that, when products of $h's$ are neglected, the Einstein equations *in vacuo* are equivalent to

$$R_{\alpha\beta} \cong -\frac{1}{2}\eta^{\lambda\mu}\partial^2_{\lambda\mu}h_{\alpha\beta} + \frac{1}{2}\partial_\alpha f_\beta + \frac{1}{2}\partial_\beta f_\alpha = 0 \tag{11.2}$$

with $\partial_\alpha := \frac{\partial}{\partial x^\alpha}$ and (indices raised with η) :

$$f_\alpha := \partial_\lambda h^\lambda_\alpha - \frac{1}{2}\partial_\alpha h^\lambda_\lambda. \tag{11.3}$$

When $f_\alpha = 0$, Equations (11.2) are wave equations for h

$$\eta^{\lambda\mu}\partial^2_{\lambda\mu}h_{\alpha\beta} = 0. \tag{11.4}$$

Equations (11.4) are a system of wave equations for the perturbation h. A solution, called **gravitational wave**, propagates with velocity 1, i.e. the speed of light.

The equations $f_\alpha = 0$ are interpreted as **polarization** conditions satisfied by the gravitational waves.

[30] Damour, T. "General Relativity Today", gr.qc 0704.0754v1, to be published in *Gravitation and Experiment*, Birkhäuser.

[31] We could also use the linearized Einstein tensor given in Chapter 1.

Equations (11.4) together with $f_\alpha = 0$ are the equations proposed by Fierz and Pauli in 1939 for a particle with rest mass zero and spin 2 in Special Relativity. It is why an eventual elementary particle associated with the gravitational field, the graviton, is said to have spin 2.

11.2 *General linear waves*

A straightforward computation using the expressions for the Ricci tensor and the Christoffel symbols gives the identity

$$R_{\alpha\beta} \equiv R^{(h)}_{\alpha\beta} + L_{\alpha\beta} \tag{11.5}$$

where

$$L_{\alpha\beta} \equiv \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta F^\lambda + g_{\beta\lambda}\partial_\alpha F^\lambda\}, \quad \text{with} \quad F^\lambda := g^{\alpha\beta}\Gamma^\lambda_{\alpha\beta}, \tag{11.6}$$

and

$$R^{(h)}_{\alpha\beta} \equiv -\frac{1}{2}g^{\lambda\mu}\partial^2_{\lambda\mu}g_{\alpha\beta} + H^{\rho\sigma\gamma\delta\lambda\mu}_{\alpha\beta}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu}, \tag{11.7}$$

where $H(g)$ is a polynomial in g and its contravariant associate[32].

Definition 11.1 *Local coordinates such that $F^\alpha = 0$ are called*[33] ***harmonic coordinates*** *or* ***wave coordinates****.*

In wave coordinates the Einstein equations

$$R^{(h)}_{\alpha\beta} = 0 \tag{11.8}$$

reduce to a quasidiagonal system of non-linear wave equations in the spacetime metric.

If we neglect products of derivatives of the metric the equations on R^{n+1} reduce to a system of wave equations for the coefficients $g_{\alpha\beta}$ considered as scalar functions,

$$g^{\lambda\mu}\partial^2_{\lambda\mu}g_{\alpha\beta} = 0. \tag{11.9}$$

Remark 11.2 *Under the conditions $F^\alpha = 0$ on the coordinates the above equations are indeed wave equations in the metric g on the scalars $g_{\alpha\beta}$ because for a scalar function u*

$$\Box_g u := \nabla^\lambda\partial_\lambda u \equiv g^{\lambda\mu}(\partial^2_{\lambda\mu}u - \Gamma^\alpha_{\lambda\mu}\partial_\alpha u). \tag{11.10}$$

We will return to wave coordinates in Chapter 6, and we will return, in a non-linear context, to the propagation of general gravitational fields.

[32] An explicit expression is given in Chapter 6.

[33] The harmonic coordinates were apparently already known by Einstein. They were used by de Donder, Lanczos, and Darmois in the 1920s.

12 High-frequency gravitational waves

The perturbation argument given in the preceding section does not justify the splitting of the equations into wave equations and polarization conditions; moreover, it has no chance to take into account non-linear effects of the gravitational field.

A powerful approximation method to study wave propagation in linear media is the WKB[34] *anzatz*, first used to study approximate solutions of the Schrödinger equation under the form

$$u(x) = a(x)e^{i\omega\phi(x)}, \tag{12.1}$$

where x is a point on spacetime, and a and ϕ are functions on spacetime. The function ϕ, called the phase of the wave, and ω, a number called frequency, are real. The number ω is assumed to be large compared with the functions a and ϕ and their derivatives with respect to x. Such a type of approximation is called in mechanics a **high-frequency wave**, or a **progressive wave** with wave fronts $\phi(x) =$ constant; the function u varies more rapidly in directions transversal to the wave fronts than on the wave fronts. The study of progressive waves, also called the "**two timing method**" has many applications in classical as well as in quantum physics.

The WKB *ansatz* is not well adapted for the study of progressive waves associated to non linear equations. Leray[35] and his collaborators introduced a more general *anzatz* which permits its use in the non linear case[36]. They replace $e^{i\omega\phi(x)}$ by a general function of x and $\omega\phi(x)$. We return to the general study of progressive gravitational waves in a later chapter.

We study in this section gravitational waves which are weak perturbations of a background in a sense we define below.

Definition 12.1 *A **weak progressive gravitational wave** on a spacetime $(V, \underline{g})$ is a C^2 Lorentzian metric g of the form*

$$g_{\alpha\beta}(x, \omega\phi(x))) \equiv \underline{g}_{\alpha\beta}(x) + \omega^{-2}v_{\alpha\beta}(x, \omega\phi(x)) \tag{12.2}$$

It is an asymptotic solution of the vacuum Einstein equations if the Ricci tensor of g satisfies an equation

$$R_{\alpha\beta} = \omega^{-2}\mathcal{R}_{\alpha\beta}(x, \omega), \tag{12.3}$$

where $\mathcal{R}(x, \omega)$, $x \in V$, $\omega \in R$ is uniformly bounded in $\omega > 0$ for some norm on V.

To compute and estimate the Ricci tensor of the metric (12.2) we use the following elementary **derivation formula**.

[34] Wentzel–Kramers–Brillouin.
[35] Garding, L., Kotake, T., and Leray, J. (1966) *Bull. Soc. Math. Fr.*, **94**, 25–48.
[36] Choquet-Bruhat, Y. (1969) *J. Math. Pures et Appl.*, **48**, 117–52.

Proposition 12.2 *The derivative with respect to x of a function of x and $\omega\phi(x)$ is given by*

$$\frac{\partial}{\partial x^\alpha} f(x, \omega\phi(x)) \equiv \{\partial_\alpha f(x,\xi) + \omega\phi_\alpha f'(x,\xi)\}_{|\xi=\omega\phi(x)}, \tag{12.4}$$

where we have set

$$\partial_\alpha f := \frac{\partial}{\partial x^\alpha} f(x,\xi) \quad \phi_\alpha := \frac{\partial \phi}{\partial x^\alpha}, \quad f' := \frac{\partial}{\partial \xi} f(x,\xi). \tag{12.5}$$

An important observation for checking the asymptotic validity for large ω (i.e. bounding F uniformly in $\omega > 0$) is the following **null average *ansatz.***

Proposition 12.3 *A function $f(.,\xi)$ with continuous derivative f' with respect to ξ can be uniformly bounded in ξ only if*

$$\lim_{T=\infty} \frac{1}{T} \int_0^T f'(.,\xi) d\xi = 0, \tag{12.6}$$

in particular f' cannot be independent of ξ.

If f' is periodic in ξ with period T, its primitive is bounded, and periodic of period T, if and only if

$$\int_0^T f'(.,\xi) d\xi = 0. \tag{12.7}$$

We deduce from the definition (12.2) of $g_{\alpha\beta}$ that its contravariant associate admits an expansion of the form

$$g^{\alpha\beta}(x, \omega\phi) = \underline{g}^{\alpha\beta} - \omega^{-2} v^{\alpha\beta}(x, \omega\phi) + \omega^{-3} g^{(3)\alpha\beta}(x,\omega) \tag{12.8}$$

where indices of v, as well as all indices in the following are raised with $\underline{g}^{\alpha\beta}$. The tensor $g^{(3)\alpha\beta}$ is uniformly bounded in $\omega > 0$ as long as it is so of $v_{\alpha\beta}$, and the determinant of g is bounded away from zero.

A consequence of the proposition 2 (derivation formula) is the following expansion of the Christoffel symbols of g

$$\Gamma^\lambda_{\alpha\beta} = \underline{\Gamma}^\lambda_{\alpha\beta} + \omega^{-1}\Gamma^{\lambda(1)}_{\alpha\beta} + \omega^{-2}\Gamma^{\lambda(2)}_{\alpha\beta} + O(\omega^{-3}), \tag{12.9}$$

with

$$\Gamma^{\lambda(1)}_{\alpha\beta} := \gamma^\lambda_{\alpha\beta} \equiv \frac{1}{2}(\phi_\alpha v'^\lambda_\beta + \phi_\beta v'^\lambda_\alpha - \phi^\lambda v'_{\alpha\beta}), \quad \text{hence} \quad \gamma^\lambda_{\alpha\lambda} := \frac{1}{2}\phi_\alpha v'^\lambda_\lambda, \tag{12.10}$$

$$\Gamma^{\lambda(2)}_{\alpha\beta} := \varphi^\lambda_{\alpha\beta} \equiv \frac{1}{2}(\underline{\nabla}_\alpha v^\lambda_\beta + \underline{\nabla}_\beta v^\lambda_\alpha - \underline{\nabla}^\lambda v_{\alpha\beta}), \quad \text{hence} \quad \varphi^\lambda_{\alpha\lambda} :\equiv \frac{1}{2}\underline{\nabla}_\alpha(v^\lambda_\lambda), \tag{12.11}$$

where indices are raised and the covariant derivative is taken with the metric $\underline{g}$; namely, for a vector u depending on x and ξ we have set

$$(\underline{\nabla}_\alpha u^\beta)(x,\omega\phi) = \left\{ \frac{\partial}{\partial x^\alpha} u^\beta(x,\xi) + \underline{\Gamma}^\beta_{\alpha\lambda}(x) u^\lambda(x,\xi) \right\}_{\xi\,=\,\omega\phi}. \tag{12.12}$$

Remark 12.4 *To compute tensorial quantities at a point one can assume* $\underline{\Gamma} = 0$ *at this point and replace in the obtained formulae first ordinary derivatives by covariant derivatives with respect to* $\underline{g}$.

Recalling that

$$R_{\alpha\beta} \equiv \partial_\lambda \Gamma^\lambda_{\alpha\beta} - \partial_\alpha \Gamma^\lambda_{\beta\lambda} + \Gamma^\lambda_{\alpha\beta}\Gamma^\mu_{\lambda\mu} - \Gamma^\lambda_{\alpha\mu}\Gamma^\mu_{\beta\lambda} \tag{12.13}$$

the previous formulas show that

$$R_{\alpha\beta}(x, \omega\phi) = R^{(0)}_{\alpha\beta}(x, \omega\phi) + \omega^{-1} R^{(1)}_{\alpha\beta}(x, \omega\phi) + \omega^{-2}\mathcal{R}_{\alpha\beta}(x, \omega). \tag{12.14}$$

with $\mathcal{R}_{\alpha\beta}$ bounded on V, uniformly for $\omega > 0$, if v is uniformly C^2 on V and the determinant of g is bounded away from zero. The metric (12.2) defines a progressive gravitational wave if

$$R^{(0)}_{\alpha\beta}(x, \omega\phi(x)) = 0, \quad \text{and} \quad R^{(1)}_{\alpha\beta}(x, \omega\phi(x)) = 0. \tag{12.15}$$

12.1 *Phase and polarizations*

A straightforward calculation gives

$$R^{(0)}_{\alpha\beta}(x, \omega\phi(x)) \equiv \underline{R}_{\alpha\beta} - \frac{1}{2}\{\underline{g}^{\lambda\mu}\phi_\lambda\phi_\mu v''_{\alpha\beta} - \phi_\beta P''_\alpha - \phi_\alpha P''_\beta\}(x, \omega\phi(x)). \tag{12.16}$$

with

$$P''_\alpha := \phi^\mu \left(v''_{\mu\alpha} - \frac{1}{2}\underline{g}_{\mu\alpha} v''^\rho_\rho \right) , \tag{12.17}$$

There are two possibilities to obtain $R^{(0)}_{\alpha\beta} = 0$

1. The phase ϕ satisfies the eikonal equation of the background spacetime

$$\underline{g}^{\lambda\mu}\phi_\lambda\phi_\mu = 0. \tag{12.18}$$

The perturbation v must then satisfy the linear system $P''_\alpha = 0$. Due to Proposition 12.3 (boundedness *ansatz*) we impose that

$$P_\alpha \equiv \phi^\mu \left(v_{\mu\alpha} - \frac{1}{2}\underline{g}_{\mu\alpha} v^\rho_\rho \right) = 0. \tag{12.19}$$

These equations are the **polarization conditions** of the wave. Remark that they are not gauge conditions but necessary conditions to be satisfied by the wave[37].

Remark 12.5 *The conditions* $\phi^\lambda\phi_\lambda = 0$ *and* $P_\alpha = 0$ *imply that*

$$v^{\lambda\mu}\phi_\lambda\phi_\mu = 0. \tag{12.20}$$

[37] Compare with the splitting of the Ricci tensor via harmonic gauge in Chapter 6 and further comments in Chapter 11.

This property is analogous to the ***exceptionality condition*** *found in the study of shocks by Lax and Boillat*[38] *and for progressive waves in general quasilinear systems in Choquet-Bruhat (1969) (see Chapter 11). It plays a role for strong high-frequency waves, and proofs of global existence of solutions (see Chapter 15).*

2. The phase ϕ is not isotropic for the background metric

$$\underline{g}^{\lambda\mu}\phi_\lambda\phi_\mu \neq 0. \tag{12.21}$$

The Equations (12.15) and Proposition 12.2 show then that $v_{\alpha\beta}$ is of the form

$$v_{\alpha\beta} = \phi_\alpha f_\beta + \phi_\beta f_\alpha. \tag{12.22}$$

Lemma 12.6 *A high-frequency wave of the form (12.21) is not significant; it can be made to vanish (up to terms bounded by* ω^{-3}*) by a change of coordinates preserving the Minkowski metric.*

Proof Consider the change of coordinates

$$x^\alpha = \tilde{x}^\alpha + \omega^{-3}h^\alpha(\tilde{x}, \omega\phi(\tilde{x})) \tag{12.23}$$

The metric (12.2) is transformed into

$$\tilde{g}_{\lambda\mu} = (\underline{g}_{\alpha\beta} + \omega^{-2}v_{\alpha\beta})\frac{\partial x^\alpha}{\partial \tilde{x}^\lambda}\frac{\partial x^\beta}{\partial \tilde{x}^\mu} \tag{12.24}$$

with

$$\frac{\partial x^\alpha}{\partial \tilde{x}^\lambda} = \delta^\alpha_\lambda + \omega^{-2}\phi_\lambda h'^\alpha + \omega^{-3}\tilde{\partial}_\lambda h^\alpha. \tag{12.25}$$

Therefore

$$\tilde{g}_{\lambda\mu} = \underline{g}_{\lambda\mu} + \omega^{-2}(v_{\lambda\mu} + \underline{g}_{\lambda\beta}\phi_\mu h'^\beta + \underline{g}_{\mu\alpha}\phi_\lambda h'^\alpha) + \omega^{-3}M_{\lambda\mu}, \tag{12.26}$$

the perturbation becomes of order ω^{-3} by appropriate choice of h if v is of the form (12.22). □

12.1.1 *Propagation equations*

We use the Remark 12.4 to compute $R^{(1)}_{\alpha\beta}$. It gives

$$R^{(1)}_{\alpha\beta} \equiv \underline{\nabla}_\lambda\gamma^\lambda_{\alpha\beta} - \underline{\nabla}_\alpha\gamma^\lambda_{\beta\lambda} + \phi_\lambda\varphi'^\lambda_{\alpha\beta} - \phi_\alpha\varphi'^\lambda_{\beta\lambda}; \tag{12.27}$$

[38] Boillat, G. (1996) in *Recent Mathematical Methods in Non Linear Wave Propagation*, Springer Lecture Notes in Maths, 1640.

that is,

$$R^{(1)}_{\alpha\beta} \equiv \frac{1}{2}\{\underline{\nabla}_\lambda(\phi_\alpha v'^{\lambda}_{\beta} + \phi_\beta v'^{\lambda}_{\alpha} - \phi^\lambda v'_{\alpha\beta}) - \underline{\nabla}_\alpha(\phi_\beta v'^{\lambda}_{\lambda})\} \tag{12.28}$$

$$+ \frac{1}{2}\{\phi_\lambda(\underline{\nabla}_\alpha v'^{\lambda}_{\beta} + \underline{\nabla}_\beta v'^{\lambda}_{\alpha} - \underline{\nabla}^\lambda v'_{\alpha\beta}) - \phi_\alpha(\underline{\nabla}_\beta v'^{\lambda}_{\lambda})\}. \tag{12.29}$$

We use the identities (recall that ϕ_α is a gradient)

$$\phi_\lambda(\underline{\nabla}_\alpha v'^{\lambda}_{\beta}) \equiv \underline{\nabla}_\alpha(\phi_\lambda v'^{\lambda}_{\beta}) - v'^{\lambda}_{\beta}\underline{\nabla}_\alpha\phi_\lambda \quad \text{and} \quad \underline{\nabla}_\alpha\phi_\lambda = \underline{\nabla}_\lambda\phi_\alpha \tag{12.30}$$

to write the equations $R^{(1)}_{\alpha\beta} = 0$ under the following form, when v satisfies the polarization conditions $\phi_\lambda v'^{\lambda}_{\alpha} = \frac{1}{2}\phi_\alpha v'^{\lambda}_{\lambda}$,

$$-\mathcal{P}_{\alpha\beta} + \frac{1}{2}\left\{\phi_\beta\left(\underline{\nabla}_\mu v'^{\mu}_{\alpha} - \frac{1}{2}\partial_\alpha v'^{\mu}_{\mu}\right) + \phi_\alpha\left(\underline{\nabla}_\mu v'^{\mu}_{\beta} - \frac{1}{2}\partial_\beta v'^{\mu}_{\mu}\right)\right\} = 0 \tag{12.31}$$

with

$$\mathcal{P}_{\alpha\beta} := \phi^\lambda\underline{\nabla}_\lambda v'_{\alpha\beta} + \frac{1}{2}\underline{\nabla}_\lambda\phi^\lambda v'_{\alpha\beta} \tag{12.32}$$

The equation $\mathcal{P}_{\alpha\beta} = 0$ is a propagation equation along the rays of the phase ϕ for the tensor v. The added terms have a form which we have shown to be not significant in the case of the progressive waves.

12.2 *Radiative coordinates*

To give a simple analysis of the propagation properties of the progressive wave we take coordinates, which we call radiative, such that the phase ϕ is constant on a coordinate surface; namely we set

$$x^0 = \phi \quad \text{hence} \quad \phi_0 = 1, \quad \phi_i = 0. \tag{12.33}$$

Taking the phase ϕ isotropic for the background $\underline{g}$ implies that

$$\underline{g}^{00} = 0, \quad \text{and} \quad \underline{g}^{0i} = \phi^i. \tag{12.34}$$

In such coordinates the significant components of the wave v are the v_{ij}.

Simple computation gives

$$v^0_0 = \phi^i v_{0i}, \quad v^0_i = \phi^j v_{ij}, \quad v^j_i = \underline{g}^{jh}v_{ih} + \phi^j v_{i0}. \tag{12.35}$$

12.2.1 *Polarization conditions*

The polarization conditions (12.19) can be written, using (12.33),

$$v^0_\alpha - \frac{1}{2}\delta^0_\alpha v^\lambda_\lambda = 0;$$

that is, respectively for $\alpha = 0$ and $\alpha = i$,

$$\underline{g}^{jh}v_{ih} = 0 \qquad \text{and} \qquad \phi^j v_{ij} = 0; \tag{12.36}$$

that is, the significant part of the tensor v at a point is traceless and orthogonal to the null ray $grad\phi$. In particular, in the classical case of a four-dimensional

spacetime, this significant part lies in a two-dimensional vector space. One says that the gravitational wave has two degrees of polarization.

12.2.2 *Propagation equations*

The equations $\mathcal{P}_{ij} = 0$ reduce to

$$\mathcal{P}_{ij} \equiv \phi^\lambda \underline{\nabla}_\lambda v'_{ij} + \frac{1}{2}\underline{\nabla}_\lambda \phi^\lambda v'_{ij} = 0 \tag{12.37}$$

they are propagation equations along the rays of the phase ϕ.

Lemma 12.7 *The propagation equations (12.37) preserve the polarization conditions (12.36)*

Proof Contraction of (12.37) with $\underline{g}^{ij}$ shows that $\underline{g}^{ij}v'ij$ propagates along the rays of ϕ.

On the other hand, contracted product with ϕ^i gives

$$\phi^\lambda \underline{\nabla}_\lambda(\phi^i v'_{ij}) + \frac{1}{2}\underline{\nabla}_\lambda \phi^\lambda(\phi^i v'_{ij}) - (\phi^\lambda \underline{\nabla}_\lambda \phi^i)v'_{ij} = 0 \tag{12.38}$$

but $\phi^\lambda \underline{\nabla}_\lambda \phi^i = \phi^\lambda \underline{\nabla}^i \phi_\lambda$, as already remarked, and is zero because ϕ_λ has constant length (zero). □

12.2.3 *Gauge conditions*

We have satisfied, modulo the polarization conditions, the significant equations $R^{(1)}_{ij} = 0$, but we have not satisfied the equations $R^{(1)}_{0\alpha} = 0$. These are linear equations for the non-significant part $v_{0\alpha}$ of the wave, once v_{ij} is known. Note that we can make $v_{0\alpha} = 0$ by a change of coordinates of the form (12.23). Such a change of coordinates introduces an additional term in ω^{-3} in the expansion of g which becomes of the form

$$g_{\alpha\beta} = \underline{g}_{\alpha\beta} + \omega^{-2}v_{\alpha\beta} + \omega^{-3}w_{\alpha\beta}. \tag{12.39}$$

The additional term $\omega^{-3}w_{\alpha\beta}$ introduces in $R^{(1)}_{\alpha\beta}$, if the phase ϕ is isotropic for $\underline{g}$, the additional term

$$\phi_\beta\left(\phi^\mu w''_{\mu\alpha} - \frac{1}{2}\phi_\alpha w''^\rho{}_\rho\right) + \phi_\alpha\left(\phi^\mu w''_{\mu\beta} - \frac{1}{2}\phi_\beta w''^\rho{}_\rho\right). \tag{12.40}$$

We see that we can solve the equations $R^{(1)}_{0\alpha}$ in radiative coordinates by choosing $v_{0\alpha} = 0$ and choosing w such that

$$\underline{\nabla}_\mu v'^\mu_\alpha - \frac{1}{2}\partial_\alpha v'^\mu_\mu + \phi^\mu w''_{\mu\alpha} - \frac{1}{2}\phi_\alpha w''^\rho{}_\rho = 0. \tag{12.41}$$

12.3 *Energy conservation*

The propagation equation 12.37 implies the conservation law

$$\underline{\nabla}_\lambda(\phi^\lambda \bar{v}'^{ij} v'_{ij}) = 0, \qquad \bar{v}'^{ij} := \underline{g}^{ih}\underline{g}^{jk} v'_{hk}. \tag{12.42}$$

One denotes by E the positive quantity[39] interpreted as the energy density of the progressive wave:

$$E := \frac{1}{4}\bar{v}'^{ij} v'_{ij} \quad \text{with} \quad \bar{v}'^{ij} := \underline{g}^{ih}\underline{g}^{jk} v'_{hk}. \tag{12.43}$$

A straightforward computation using the relations (12.35) shows that E is a scalar function given in arbitrary coordinates by

$$E = \frac{1}{4}\left\{ v'^{\alpha\beta} v'_{\alpha\beta} - \frac{1}{2}(v'^{\alpha}_{\alpha})^2 \right\}. \tag{12.44}$$

In arbitrary coordinates the conservation law (12.42) reads

$$\underline{\nabla}_\lambda(\phi^\lambda E) = 0, \tag{12.45}$$

it is interpreted as the conservation by propagation through the flow of the rays ϕ_λ of the energy of the progressive wave.

13 Coupled electromagnetic and gravitational waves

We look for an electrovac Einsteinian spacetime which exhibits the coupling of gravitational and electromagnetic progressive wave. If these waves are of the same order, the electromagnetic field is of the form

$$F_{\alpha\beta}(x) := \underline{F}_{\alpha\beta}(x) + \omega^{-1} H_{\alpha\beta}(x, \omega\phi(x)), \tag{13.1}$$

where $\underline{F}$ is a background field and H the progressive wave.

The Maxwell tensor τ admits the expansion

$$\tau_{\alpha\beta} = \underline{\tau}_{\alpha\beta} + \omega^{-1}\{\underline{F}_\alpha{}^\lambda H_{\beta\lambda} + \underline{F}_{\beta\lambda} H_\alpha{}^\lambda\} + O(\omega^{-2}) \tag{13.2}$$

For simplicity we consider the most physically interesting case, a spacetime of dimension 4. The Maxwell tensor has then a zero trace and the electrovac Einstein equations are

$$R_{\alpha\beta} - \tau_{\alpha\beta} = 0, \quad \tau_{\alpha\beta} \equiv F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}. \tag{13.3}$$

The pair (v, H) is a coupled asymptotic electrovac progressive wave on the manifold V with background metric and electromagnetic field $(\underline{g},\underline{F})$ if (g, F) is such that

$$(R_{\alpha\beta} - \tau_{\alpha\beta})(x, \omega\phi(x)) = \omega^{-2} M(x, \omega) \tag{13.4}$$

with M bounded on V, uniformly for $\omega > 0$.

[39] The factor 1/4 is introduced for coherence with other types of energy; see next section.

13.1 *Phase and polarizations*

The terms of order zero in ω in the expansion of $R_{\alpha\beta} - \tau_{\alpha\beta}$ are

$$R^{(0)}_{\alpha\beta} - \underline{\tau}_{\alpha\beta} \tag{13.5}$$

with $R^{(0)}_{\alpha\beta}$ given by (12.16). as in the vacuum case. The averaging *ansatz* (Proposition 12.3) implies therefore that the background ($\underline{g}$,$\underline{F}$) satisfies the Einstein equations with source $\underline{F}$

$$\underline{R}_{\alpha\beta} = \underline{\tau}_{\alpha\beta}, \tag{13.6}$$

while the phase must still be isotropic for the gravitational wave to be significant; the polarization conditions of this wave are unchanged.

The Maxwell equations admit expansions in powers of ω with coefficients[40] of order zero given by

$$(dF)^{(0)}_{\alpha\beta\gamma} = (d\underline{F})_{\alpha\beta\gamma} + \phi_\alpha H'_{\beta\gamma} + \phi_\gamma H'_{\alpha\beta} + \phi'_\beta H_{\gamma\alpha}, \tag{13.7}$$

$$(\nabla_\alpha F^\alpha{}_\beta)^{(0)} = \underline{\nabla}_\alpha \underline{F}^\alpha{}_\beta + \phi^\alpha H'{}_{\alpha\beta}. \tag{13.8}$$

They are satisfied at order zero if the background is an electrovac Einsteinian spacetime and the wave H is such that

$$\phi_\alpha H'_{\beta\gamma} + \phi_\gamma H'_{\alpha\beta} + \phi'_\beta H_{\gamma\alpha} \qquad \text{and} \qquad \phi^\alpha H'_{\alpha\beta} = 0. \tag{13.9}$$

Contraction of the first equation with ϕ^α and use of the second equation show that these equations imply $H'_{\beta\gamma} = 0$; that is, the vanishing of the wave H except if $\phi^\alpha\phi_\alpha = 0$, i.e. if the phase ϕ is isotropic. In radiative coordinates we then have

$$H'_{ij} = 0, \qquad \phi^i H'_{i0} = 0. \tag{13.10}$$

These linear equations are the **polarization conditions** of the electromagnetic wave. They read in arbitrary coordinates, since H is antisymmetric

$$H_{\alpha\beta} = a_\alpha\phi_\beta - a_\beta\phi_\alpha, \quad \text{and} \quad \phi^\alpha a_\alpha = 0, \tag{13.11}$$

a is a vector orthogonal to the ray ϕ^α.

13.2 *Propagation equations*

13.2.1 *Gravitational wave*

The conditions of order 1 to be satisfied by the Einstein equations are

$$R^{(1)}_{\alpha\beta} - \tau^{(1)}_{\alpha\beta} = 0. \tag{13.12}$$

[40] Notations as in the previous section, indices raised with $\underline{g}$.

They read in radiative coordinates, modulo the polarization conditions, as propagation equations with the same principal part as in the vacuum case, but now with a source, the electromagnetic wave

$$\phi^\lambda \underline{\nabla}_\lambda v'_{ij} + \frac{1}{2}\underline{\nabla}_\lambda \phi^\lambda v'_{ij} + \underline{g}_{ij}\underline{F}^{0h}a_h - \underline{F}^0_i a_j - \underline{F}^0_j a_i = 0. \tag{13.13}$$

Lemma 13.1 *The source term in the gravitational wave propagation equation (13.13) vanishes for a non-zero electromagnetic wave if and only if $\phi_\lambda \underline{F}^{\lambda\mu}$ is colinear to ϕ^μ.*

Proof One takes at a point a space frame such that $\underline{g}_{ij} = \delta_{ij}$ and $\underline{F}^{0h} = A\delta^h_1$. It is then easy to check that the system written in radiative coordinates

$$\underline{g}_{ij}\underline{F}^{0h}a_h - \underline{F}^0{}_i a_j - \underline{F}^0{}_j a_i = 0 \tag{13.14}$$

implies $a_i = 0$ if $i \neq 1$, and $a_1 = 0$ if $A \neq 0$, i.e. $\underline{F}^{0h} = 0$; that is, in arbitrary coordinates

$$\phi_\lambda \underline{F}^{\lambda\mu} = k\phi^\mu. \tag{13.15}$$

□

13.2.2 *Electromagnetic wave*

Maxwell equations are satisfied at order 1 in ω if

$$(dF)^{(1)}_{\alpha\beta\gamma} \equiv (dH)_{\alpha\beta\gamma} \equiv \partial_\alpha H_{\beta\gamma} + \partial_\gamma H_{\alpha\beta} + \partial_\beta H_{\gamma\alpha} = 0 \tag{13.16}$$

and

$$(\nabla_\alpha F^\alpha{}_\beta)^{(1)} \equiv \underline{\nabla}_\alpha H^\alpha{}_\beta + \Gamma^{\alpha(1)}_{\alpha\lambda}\underline{F}^\lambda{}_\beta - \Gamma^{\lambda(1)}_{\alpha\beta}\underline{F}^\alpha{}_\lambda = 0, \tag{13.17}$$

with

$$\underline{\nabla}_\alpha H^\alpha{}_\beta \equiv \phi_\beta \underline{\nabla}^\alpha a_\alpha + a_\alpha \underline{\nabla}^\alpha \phi_\beta - \phi_\alpha \underline{\nabla}^\alpha a_\beta - a_\beta \underline{\nabla}^\alpha \phi_\alpha \tag{13.18}$$

We have using a previous remark, an elementary identity, and the condition (13.11)

$$a_\alpha \underline{\nabla}^\alpha \phi_\beta \equiv a_\alpha \underline{\nabla}_\beta \phi^\alpha \equiv \underline{\nabla}_\beta(a_\alpha \phi^\alpha) - \phi^\alpha \underline{\nabla}_\beta a_\alpha = -\phi^\alpha \underline{\nabla}_\beta a_\alpha. \tag{13.19}$$

We take, as in the previous section, radiative coordinates. Then $H_{ij} = 0$ and $H_{i0} = a_i$, and Equation (13.16) reads

$$\partial_i a_j - \partial_j a_i = 0. \tag{13.20}$$

Therefore we have, modulo this equation,

$$\underline{\nabla}_\alpha H^\alpha{}_i = -2\phi^\alpha \underline{\nabla}_\alpha a_i - a_i \underline{\nabla}_\alpha \phi^\alpha. \tag{13.21}$$

The expression (12.10) of $\Gamma^{\lambda(1)}_{\alpha\beta}$ gives for a polarized gravitational wave $\Gamma^{\alpha(1)}_{\alpha\lambda} = 0$ and, in radiative coordinates

$$\Gamma^{\lambda(1)}_{\alpha i}\underline{F}^\alpha{}_\lambda = \underline{F}^{0j} v'_{ij}. \tag{13.22}$$

Equation (13.17) reduces therefore to a linear propagation for a with source $\underline{F}^{0j}v'_{ij}$.

$$2\phi^{\alpha}\underline{\nabla}_{\alpha}a_i + a_i\underline{\nabla}_{\alpha}\phi^{\alpha} + \underline{F}^{0j}v'_{ij} = 0. \tag{13.23}$$

We deduce from this equation the following intrinsic result.

Lemma 13.2 *The source term in the electromagnetic wave propagation equation vanishes for a non-zero gravitational wave if and only if $\phi_{\lambda}\underline{F}^{\lambda\mu}$ is colinear to ϕ^{μ}.*

13.2.3 *Energy conservation*

We remark that, modulo the polarization condition $v'^{ij}\underline{g}_{ij} = 0$, the propagation equations (13.13) and (13.23) imply the conservation law

$$\underline{\nabla}_{\alpha}\left\{\phi^{\alpha}\left(a^i a_i + \frac{1}{4}v'^{ij}v'_{ij}\right)\right\} = 0. \tag{13.24}$$

It is interpreted as the conservation by propagation of the total energy of the coupled gravitational and electromagnetic waves, with energy density

$$a^i a_i + \frac{1}{4}v'^{ij}v'_{ij}. \tag{13.25}$$

In arbitrary coordinates this conservation law reads

$$\underline{\nabla}_{\alpha}(\phi^{\alpha}\mathcal{E}) = 0, \tag{13.26}$$

with

$$\mathcal{E} \equiv \mathcal{E}_{grav} + \mathcal{E}_{e\cdot m} := \frac{1}{4}\left\{v'^{\alpha\beta}v'_{\alpha\beta} - \frac{1}{2}\left(v'^{\alpha}_{\alpha}\right)^2\right\} + \frac{1}{2}H^{\alpha\beta}H_{\alpha\beta}. \tag{13.27}$$

13.2.4 *Conclusion*

We state as a theorem the results that we have obtained.

Theorem 13.3 *In an electrovac Einsteinian background $(\underline{g}, \underline{F})$, a progressive gravitational [respectively electromagnetic] wave with isotropic phace ϕ generates a progressive electromagnetic [respectively gravitational] wave of the same phase if $\phi_{\lambda}\underline{F}^{\lambda}{}_{\mu}$ is not collinear to ϕ_{μ}.*

The total energy of the waves is conserved by propagation.

IV

SCHWARZSCHILD SPACETIME AND BLACK HOLES

1 Introduction

Soon after the publication of Einstein's equations, an exact solution was constructed by Schwarzschild which could model the gravitational field outside a spherically symmetric isolated body, for instance the Sun.

By Einstein's theory the paths of the planets, considered as test particles, are timelike geodesics in the Schwarzschild spacetime. These geodesics were found to correspond to the Kepler ellipses of Newton's gravity in first approximation, and (a remarkable fact in support of Einstein's theory), they accounted for the tiny advance of the perihelion of Mercury which could not be explained by Newton's theory and perturbation by other planets. In addition the observation of the new effects predicted by General Relativity – time delay and deflection of light rays – could be proved to be in agreement with their theoretical values in a Schwarzschild spacetime.

Finally the Schwarzschild solution made apparent a new phenomenon – black holes – strange objects which have no analogue in Newtonian mechanics in spite of Mitchell's and Laplace's early remarks that light could be captured by a massive star. Black holes (see Section 11) play a major role in astronomy and cosmology today.

2 Spherically symmetric spacetimes

We now state the following natural and elementary definitions.

Definition 2.1 *A three-dimensional Riemannian manifold $(M, \bar{g})$ is said to be spherically symmetric if*

1. *The manifold M is represented by one chart (U, Φ) with $\Phi(U) = R^3$, or the exterior of a ball B of R^3 centred at some point O. We denote by ρ, θ, ϕ spherical (pseudo-) coordinates in $\Phi(U)$, linked to canonical coordinates x, y, z of R^3 by the usual relations*

$$x = \rho sin\theta sin\phi, \quad y = \rho sin\theta cos\phi, \quad z = \rho cos\theta \tag{2.1}$$

2. *In $\Phi(U)$, given by $\rho \geq \rho_0 \geq 0, 0 \leq \theta < \pi, 0 \leq \varphi < 2\pi$, $\bar{g}$ is represented by a metric of the form*

$$e^{h(\rho)} d\rho^2 + f^2(\rho)(d\theta^2 + sin^2\theta d\varphi^2). \tag{2.2}$$

The interpretation is that $\Phi(U)$ is foliated by metric 2-spheres $\rho =$ constant, centred at O; their areas in the metric (2.2) are $4\pi f^2$. The metric (2.2) is the general form of a metric invariant by rotations in R^3, centred at O. It is defined on the whole of R^3 if the ball B is empty.

The choice of the coordinate r given by $r = f(\rho)$ is called the standard choice, it is admissible as long as f is a monotonically increasing function of ρ. The number $r_0 = f(\rho_0)$ corresponds to the ball B.

Remark 2.2 *$r = 0$ represents, by definition, a single point O. The vanishing of $f^2(0)$ does not imply a singularity in the metric, but reflects the fact that spherical "coordinates" are not admissible coordinates at $\rho = 0$.*

Definition 2.3 *Consider a spacetime (V, g) with V contained in the product $R^3 \times R$, a point of V being labelled (x, t). Suppose the subsets M_t of constant t are spacelike submanifolds, we denote by $\bar{g}_t$ the Riemannian metric induced by g on M_t. The trajectories of the vectors $\partial/\partial t$ are supposed timelike. The spacetime is said to be spherically symmetric if*

1. *Each manifold M_t has a representation as the exterior $R^3 - B_t$ of a ball B_t of R^3 centred at the origin O. Each manifold $(M_t, \bar{g}_t)$ is spherically symmetric. In $R^3 - B_t$ the metric $\bar{g}_t$ reads in standard coordinates*

$$\bar{g}_t = e^{\lambda(r,t)}dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2). \tag{2.3}$$

2. *For each t, the g length and the representative of the projection on M_t of the vector $\partial/\partial t$ tangent to the timeline[1], are both invariant under the rotation group defined above.*

Lemma 2.4 *A spherically symmetric space time (V, g) admits a metric of the form:*

$$g = -e^{\nu}dt^2 + e^{\lambda}dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2) \tag{2.4}$$

where λ and ν are functions of t and r only.

Proof A scalar on $R^3 - B$ invariant under the given rotation group is necessarily a function of r and t alone. A vector field invariant under this group is tangent to the radial lines (lines where only the r coordinate varies) and its magnitude depends only on r for each t. Therefore the given definition implies that:

$$g = -a^2(r,t)dt^2 + 2b(r,t)dtdr + e^{\lambda(r,t)}dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2) \tag{2.5}$$

We can eliminate the diagonal term in $dtdr$ by computing an integrating factor for the 2-form in two variables $\omega := adt - bdr$, i.e. a function of t and r, which we denote $e^{-2\nu}$, such that its product by the 2-form is the differential of a function τ:

$$e^{-2\nu}\omega \equiv e^{-2\nu}(t,r)(adt - bdr) \equiv d\tau$$

[1] Equivalently, the lapse and the shift of the slicing (see Chapter 6).

We take τ as a new time coordinate, keeping r, θ, φ as space pseudo-coordinates. Writing the metric (2.5) with this new coordinate, and renaming it t, gives the formula (2.4). □

Remark 2.5 *If the point with coordinate $r = 0$ belongs to the manifolds M_t it describes a timelike line called the central world line.*

3 Schwarzschild metric

We will prove the following theorem.

Theorem 3.1 *A smooth spherically symmetric metric is a solution of the vacuum Einstein equations if and only if it is the Schwarzschild metric which reads in standard coordinates*

$$g_{Schw} = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2). \quad (3.1)$$

Proof We set $t = x^0$, $r = x^1$, $\theta = x^2, \varphi = x^3$ and denote by a prime the derivative with respect to r. We find that the only non-zero Christoffel symbols of the spherically symmetric metric (2.4) are

$$\Gamma^0_{00} = \frac{1}{2}\partial_t\nu, \quad \Gamma^1_{00} = \frac{1}{2}e^{\nu-\lambda}\nu', \quad \Gamma^0_{01} = \frac{\nu'}{2}, \quad \Gamma^0_{11} = \frac{\partial_t\lambda}{2}e^{\lambda-\nu}, \quad \Gamma^1_{01} = \frac{\partial_t\lambda}{2} \quad (3.2)$$

$$\Gamma^1_{11} = \frac{\lambda'}{2}, \quad \Gamma^1_{22} = -re^{-\lambda}, \quad \Gamma^2_{12} = \Gamma^3_{13} = r^{-1}, \quad \Gamma^1_{33} = -rsin^2\theta e^{-\lambda} \quad (3.3)$$

$$\Gamma^2_{33} = -sin\theta cos\theta, \quad \Gamma^3_{23} = cot\theta. \quad (3.4)$$

The components of the Ricci tensor $R_{20} = R_{30} = R_{23} = 0$,

$$R_{10} \equiv r^{-1}\partial_t\lambda, \quad R_{22} \equiv -e^{-\lambda}\left(1 + \frac{r}{2}(\nu' - \lambda')\right) + 1, \quad R_{33} \equiv sin^2\theta R_{22}. \quad (3.5)$$

We deduce from these identities that for a solution of the vacuum Einstein equations, $R_{\alpha\beta} = 0$, λ is independent of t. Therefore ν' is also independent of t, $\partial_t\nu' = 0$, and ν is of the form

$$\nu(t, r) = \nu(r) + f(t); \text{set } \tau := \int e^{\frac{1}{2}f(t)^2}dt \quad (3.6)$$

then $e^v dt^2 = e^{v(r)}dt^2$, we rename τ by t. The other non-identically zero components of the Ricci tensor reduce then to

$$R_{00} \equiv e^{\nu-\lambda}\left\{\frac{\nu''}{2} + \frac{\nu'^2}{4} - \frac{\nu'\lambda'}{4} + \frac{\nu'}{r}\right\} \quad (3.7)$$

$$R_{11} \equiv -\frac{\nu''}{2} - \frac{\nu'^2}{4} + \frac{\lambda'\nu'}{4} + \frac{\lambda'}{r} \quad (3.8)$$

The vacuum Einstein equations imply therefore

$$r(e^{\lambda-\nu}R_{00} + R_{11}) \equiv \nu' + \lambda' = 0. \tag{3.9}$$

Modulo this relation, the equations $R_{22} = 0$ and $R_{33} = 0$ reduce to

$$-e^{-\lambda}(1 - r\lambda') + 1 = 0;$$

that is

$$(e^{-\lambda})' + \frac{e^{-\lambda}}{r} = \frac{1}{r}. \tag{3.10}$$

The general solution of this linear equation for $e^{-\lambda}$ is

$$e^{-\lambda} = 1 + \frac{A}{r}, \quad \text{hence} \quad e^{\nu} = B\left(1 + \frac{A}{r}\right) \tag{3.11}$$

with A and B arbitrary constants. The constant B can be made equal to 1 by a rescaling of t. The constant A is denoted by $-2m$, and we will see that the Newtonian approximation corresponding to the Schwarzschild metric then coincides with Newton's gravity. We have thus obtained the metric (3.1).

To show that this metric satisfies the full vacuum Einstein equations we must check that it satisfies also the equation $R_{00} = 0$ and $R_{11} = 0$; this can be done by using Equations (3.9) together with (3.7), (3.8), and (3.11) to show that $e^{\lambda-\nu}R_{00} - R_{11} = 0$. □

Exercise. *Prove this last statement by using the Bianchi identities.*

Remark 3.2 *We have supposed in this theorem that the areas of the orbits of the symmetry group are monotonically increasing along their orthogonal trajectories.*

In the course of the proof we have obtained the **Birkhoff theorem:**

Theorem 3.3 *(Birkhoff) A smooth spherically symmetric metric solution of the vacuum Einstein equations is necessarily static.*

4 Other coordinates

In some problems it is useful to use alternative, non-standard, coordinates for the Schwarzschild metric.

4.1 *Isotropic coordinates*

One defines new coordinates X, Y, Z, called isotropic, on $R^3 \times R$ which are related to the standard r, θ, ϕ (i.e. x, y, z) by setting

$$r := R\left(1 + \frac{m}{2R}\right)^2, \tag{4.1}$$

$$X := Rr^{-1}x, \quad Y := Rr^{-1}y, \quad Z := Rr^{-1}z. \tag{4.2}$$

In terms of these coordinates the metric reads, with $R^2 = X^2 + Y^2 + Z^2$,

$$g_{Schw} = -\left(\frac{2R-m}{2R+m}\right)^2 dt^2 + \left(1+\frac{m}{2R}\right)^4 (dX^2 + dY^2 + dZ^2) \tag{4.3}$$

It is clear from this expression that the spaces $t =$ constant of the Schwarzschild metric are conformal to the Euclidean space.

Remark 4.1 *r is a monotonically increasing function of R from $2m$ to infinity when R increases form $m/2$ to infinity, r and R are equivalent for large r.*

4.2 *Wave coordinates*

Theorem 4.2 *The $3+1$ metric[2], defined for $\bar{r} > m$,*

$$-\frac{\bar{r}-m}{\bar{r}+m}dt^2 + \frac{\bar{r}+m}{\bar{r}-m}d\bar{r}^2 + (\bar{r}+m)^2(d\theta^2 + sin^2\theta d\phi^2) \tag{4.4}$$

is isometric to the Schwarzschild metric by the mapping $r = \bar{r} + m$. The corresponding Cartesian coordinates t and x^i defined by

$$x^1 = \bar{r}sin\theta sin\phi, \quad x^2 = \bar{r}sin\theta cos\phi, \quad x^3 = \bar{r}cos\theta,$$

are wave coordinates.

Proof We write an arbitrary spherically symmetric, static, metric on $R^3 \times R$ under the form

$$-A^2dt^2 + B^2d\bar{r}^2 + r^2(d\theta^2 + sin^2\theta d\phi^2), \tag{4.5}$$

where θ, ϕ are spherical coordinates on S^2. The coefficients A, B, and r are functions of $\bar{r}$ only. The coordinate t is obviously a wave coordinate. We look for a function $r = f(\bar{r})$ such that the coordinates x^i defined by the above relations are wave coordinates, i.e. such that the functions x^i satisfy the wave equation. For an arbitrary function ψ the wave equation in the metric (4.4) reads

$$-\frac{1}{A^2}\frac{\partial^2\psi}{\partial t^2} + \frac{1}{r^2}\left\{\frac{1}{AB}\frac{\partial}{\partial\bar{r}}\left(AB^{-1}r^2\frac{\partial\psi}{\partial\bar{r}}\right) + \Delta^*\psi\right\} \tag{4.6}$$

where Δ^* is the Laplacian on the sphere S^2. If x^i is one of the functions given above, then

$$\frac{\partial x^i}{\partial\bar{r}} = \frac{x^i}{\bar{r}}, \quad \frac{\partial^2 x^i}{\partial\bar{r}^2} = 0 \text{ and } \Delta^* x^i = -2x^i. \tag{4.7}$$

Hence the condition that the x^i be wave coordinates reduces to

$$\frac{1}{AB}\frac{d}{d\bar{r}}(AB^{-1}r^2) - 2\bar{r} = 0. \tag{4.8}$$

[2] Note that while the standard form of the Schwarzschild metric generalizes to higher values of n, its expression in wave coordinates does not (see section 14).

In the case of the Schwarzschild metric we have for $r > 2m$

$$A \equiv \sqrt{1 - \frac{2m}{r}}, \quad B \equiv A^{-1}\frac{dr}{d\bar{r}}, \tag{4.9}$$

hence the condition of harmonicity is

$$\frac{d}{dr}\left(\frac{d\bar{r}}{dr}A^2r^2\right) - 2\bar{r} \equiv \frac{d}{dr}\left(\frac{d\bar{r}}{dr}(r^2 - 2mr)\right) - 2\bar{r} = 0. \tag{4.10}$$

One sets

$$r = m(1 + z) \tag{4.11}$$

The equation becomes the Legendre equation

$$\frac{d}{dz}\left[(z^2 - 1)\frac{d\bar{r}}{dz}\right] - 2z = 0, \quad \text{for } z > 1. \tag{4.12}$$

This linear second-order differential equation is a Legendre equation whose general solution is, with C_1 and C_2 arbitrary constants

$$C_1 z + C_2\left(\frac{z}{2}ln\frac{z+1}{z-1} - 1\right) \tag{4.13}$$

To preserve the asymptotically Euclidean character of the space metric, we take

$$\bar{r} = mz, \quad \text{i.e.} \quad r = \bar{r} + m. \tag{4.14}$$

Using this value of r one may check that the Schwarzschild metric in the wave coordinates x^i defined above is the given metric. □

4.3 *Painlevé–Gullstrand-like coordinates*

The following stationary but non-static form of the Schwarzschild metric is actually used in numerical computations[3]

$$-\left(1 - \frac{2m}{r}\right)dt^2 + dx^2 + dy^2 + dz^2 + \frac{2}{r}\sqrt{\frac{2m}{r}}(xdx + ydy + zdz)dt. \tag{4.15}$$

4.4 *Regge–Wheeler coordinates*

One sets, in the region $r > 2m$,

$$\rho = r + 2m\log(r - 2m) \tag{4.16}$$

The metric takes then a form in which the timelike sections $\theta =$ constant, $\phi =$ constant are conformal to the two-dimensional Minkowski space

$$g_{Schw} = \left(1 - \frac{2m}{r}\right)(-dt^2 + d\rho^2) + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{4.17}$$

[3] Pretorius, F. (2004) gr-qc /0407110 v2.

5 Schwarzschild spacetime

If $m = 0$ the Schwarzschild metric reduces to the Minkowski metric; if $m \neq 0$ it is singular for $r = 0$.

The sign of the constant $m \neq 0$ is very important in determining the properties of the Schwarzschild metric.

If $m < 0$, the Schwarzschild metric defines a spherically symmetric spacetime on the whole of $R^3 \times R$, except at $\{0\} \times R$ where the metric is singular. At present the Schwarzschild spacetime with $m < 0$ has no physical interpretation. It can be shown that there exists no asymptotically Euclidean Einsteinian spacetime $(R^3 \times R, g)$ with sources of positive energy which coincides for $r > a \geq 0$ with a Schwarzschild metric with $m < 0$ (cf. a more general theorem in the article on the positivity of mass).

If $m > 0$, then

- For r large with respect to m the Schwarzschild metric coincides with the Newtonian approximation, since $2m/r$ is the Newtonian potential outside a spherical body of mass m centred at $r = 0$.
- Since $1-(2m/r)$ vanishes for $r = 2m$, the metric appears to be singular there. Across $r = 2m$ the timelike and spacelike character of the coordinates t and r are interchanged. We will return later to discuss this apparent "Schwarzschild singularity".

We use the name **Schwarzschild spacetime** to refer to a manifold product of $R^3 \cap \{r > a > 2m\} \cong S^2 \times R_+$ by R, which is endowed with a Schwarzschild metric. It is interpreted as the spacetime modelling the exterior of a spherically symmetric body of radial standard coordinate a and relativistic mass m. If a is identified with the radius of the body in classical length units and m with its classical mass estimated in units of length ($2m$ is then called its **Schwarzschild radius**), the solar planets and the usual stars are such that $a > 2m$. For the Sun it holds that $2m_{sun} = 2.96$ km and for the earth $2m_{earth} = 8.8$ mm.

It is possible to construct the so-called interior Schwarzschild solutions which are spherically symmetric spacetimes satisfying the Einstein equations with sources of positive energy, are smooth on a manifold $S^2 \times 0 \leq r < a$ and constitute, together with an exterior Schwarzschild spacetime, a complete admissible Einsteinian spacetime. The known interior explicit solutions are, however, constructed through equations of state which are not physically realistic and are of interest only for rough estimates.

6 The motion of the planets and perihelion precession

6.1 *Equations*

The trajectories of bodies of small size and mass in a spherically symmetric gravitational field, for instance the trajectories of the solar planets, are timelike geodesics of this field, i.e. of the Schwarzschild metric.

We denote by ds the element of proper time on a timelike curve; i.e. with our signature convention,

$$ds^2 \equiv -g_{\alpha\beta}dx^\alpha dx^\beta. \tag{6.1}$$

As in the Newtonian case we find that the orbits remain in a "plane" of R^3 by considering first the equation (cf. the expressions of the Christoffel symbols)

$$\frac{d^2\theta}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds} - sin\theta cos\theta\left(\frac{d\varphi}{ds}\right)^2 = 0. \tag{6.2}$$

One chooses the coordinate θ such that at some initial instant s_0 one has for the considered motion of the planet $\theta(s_0) = \frac{\pi}{2}$ and $\frac{d\theta}{ds}(s_0) = 0$. The equation satisfied by θ implies then that the orbit remains in the "plane" $\theta{=}\frac{\pi}{2}$. With this choice of the coordinate θ the equation for φ reduces to

$$\frac{d^2\varphi}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\varphi}{ds} = 0, \tag{6.3}$$

which integrates to an analogue of the Newtonian area law, with ℓ some constant such that

$$r^2\frac{d\varphi}{ds} = \ell. \tag{6.4}$$

Remark 6.1 *Equation (6.4) is a consequence of the invariance of the metric under rotation. The constant ℓ can be interpreted to be the angular momentum per unit mass, as seen at large distance.*

The equation

$$\frac{d^2t}{ds^2} + \frac{d\nu}{dr}\frac{dr}{ds}\frac{dt}{ds}, \quad \text{with} \quad \nu \equiv log\left(1 - \frac{2m}{r}\right) \tag{6.5}$$

integrates to ("energy" integral due to t translation invariance)

$$\left(1 - \frac{2m}{r}\right)\frac{dt}{ds} = E, \quad E \text{ a constant.} \tag{6.6}$$

The remaining geodesic equation is

$$\frac{d^2r}{ds^2} + \frac{1}{2}\frac{d\lambda}{dr}\left(\frac{dr}{ds}\right)^2 = e^{-\lambda}\left(\frac{d\varphi}{ds}\right)^2 + \frac{e^{\nu-\lambda}}{2}\frac{d\nu}{ds}\left(\frac{dt}{ds}\right)^2. \tag{6.7}$$

Using the expressions of λ and ν together with the previous integrals (6.4) and (6.6), we see that, when $\ell \neq 0$, this equation reduces to:

$$\frac{d^2u}{d\varphi^2} + u = \frac{m}{\ell^2} + 3mu^2, \quad \text{with} \quad u \equiv \frac{1}{r}. \tag{6.8}$$

This equation is formally the same as the linear equation in u of Newtonian mechanics, except for the addition of the non-linear term $3mu^2$. Therefore it holds that

$$u = u_{Newton} + v, \tag{6.9}$$

where v satisfies the equation

$$\frac{d^2v}{d\varphi^2} + v = 3mu^2. \tag{6.10}$$

As is well known the solutions of the Newtonian equation (for $\ell \neq 0$) are the conics

$$\frac{1}{r} = m\ell^{-2}(1 + e\ cos\varphi) := u_{Newton} \tag{6.11}$$

where e is the eccentricity and where the longitude of the perihelion, a constant in this Newtonian case, has been taken to be equal to zero. The constant e depends on the initial position and velocity of the planet. It is not far from 0 for the solar planets; their orbits are almost circular.

For the solar planets r is large with respect to ℓ; hence $3mu^2$ can be considered as a small correction to $m\ell^{-2}$, and the correction v will be small with respect to u_{Newton}. An approximate solution of (6.10), a small correction to the Newtonian expression, is obtained by replacing (6.10) by the linear equation

$$\frac{d^2v}{d\varphi^2} + v = 3mu^2_{Newton} \equiv 3m^3\ell^{-4}(1 + 2e\ cos\varphi + e^2\ cos^2\varphi). \tag{6.12}$$

This is the differential equation for a forced oscillation. The general solution is the sum of the general solution of the associated homogeneous equation, i.e. an arbitrary periodic function of period 2π in φ, and a particular solution, for instance:

$$3m^3\ell^{-4}\left(1 + e\varphi\ sin\varphi + e^2\left(\frac{1}{2} - \frac{1}{6}cos2\varphi\right)\right) \tag{6.13}$$

The only term which will give a significant contribution to the correction is the "secular" (non-periodic) term in φ; that is, we consider the Einsteinian approximation:

$$\begin{aligned} u_{Einstein} &\sim u_{Newton} + 3m^2\ell^{-4}e\varphi\ sin\varphi \\ &= m\ell^{-2}(1 + ecos\varphi + e3m^2\ell^{-2}\varphi sin\varphi). \end{aligned} \tag{6.14}$$

In geometric units $m^2\ell^{-2}$ is small, therefore $3m^2\ell^{-2}\varphi$ is equivalent to $sin(3m^2\ell^{-2}\varphi)$ and $u_{Einstein}$ is equivalent to

$$u_{Einstein} \sim m\ell^{-2}(1 + ecos((1 - 3m^2\ell^{-2})\varphi)) \tag{6.15}$$

The orbit is no longer a closed curve (except if it is circular, i.e. if $e = 0$), because $u_{Einstein}$ does not have period 2π in φ. It is only approximately an ellipse; the closest point to the centre (perihelion), attained for $\varphi = 0$, is attained successively after each increase of φ by $2\pi(1 - 3m^2\ell^{-2})^{-1} \sim 2\pi + 6\pi m^2\ell^{-2}$. The additional $6\pi m^2\ell^{-2}$ is the famous Einsteinian perihelion precession.

Remark 6.2 *Elementary calculus gives a perihelion precession*

$$\frac{6\pi m}{a(1-e^2)}$$

where a is the semi-major axis of the Newtonian ellipse. The precession of the perihelion is greatest for eccentricities approaching 1.

6.2 *Results of observations*

It had long been observed by astronomers that the orbits of the solar planets are not exact ellipses, but slowly rotating ones. This phenomenon was interpreted in Newton's theory as a result of the influence of other planets. The perihelion precession of most planets could thus be accounted for, except for that of the nearest to the Sun, Mercury, for which a precession of $42''$ per century (over the $5.600''$ observed) remained unexplained. The Einsteinian correction (1916) just filled the gap, which was a remarkable success for the new theory. The not exactly symmetric shape of the Sun could also play a role but is difficult to estimate (in the case of satellites orbiting around the Earth, the irregularities of their motions are used to determine the shape of the Earth).

In 1974 a pulsar (rotating neutron star, PSR 1913+16) was observed by Hulse and Taylor, orbiting in the strong gravitational field of an object whose nature is unknown (white dwarf, neutron star, or black hole). Its orbit shows a precession of about $4.2°$ per year, which is about 271 times the total precession of Mercury. It is believed to be an Einsteinian effect: since the masses of the orbiting objects are not known the precession cannot be used directly to test General Relativity; it is used instead to estimate these masses. There is an excellent agreement between the observed rates of change of the orbital period of the pulsar and the one predicted[4] from gravitational radiation using these masses and the orbital parameters.

6.3 *Escape velocity*

As an exercise in physical interpretations in General Relativity, we compute the radial velocity with respect to an observer at rest in the Schwarzschild metric which must be applied to a test object for it to escape the gravitational attraction.

Let r_0 be the r coordinate of the static observer in a Schwarzschild spacetime with mass $m < \frac{r_0}{2}$. Denote by $\dot{r}_0 = \frac{dr}{ds}(0)$ the proper time initial velocity, supposed to be radial, of the launched rocket at parameter time $t = 0$. This rocket,

[4] See Damour, T. and Taylor, J. H. (1991) *Astrophys. J.*, **366**, 501.

supposed to be motorless, follows a radial geodesic curve and hence satisfies, by (6.6), the equation:

$$\left(1-\frac{2m}{r}\right)\dot{t}=E=\left(1-\frac{2m}{r_0}\right)\dot{t}_0. \tag{6.16}$$

We also have, as a result of the definition of ds,

$$1=\left(1-\frac{2m}{r}\right)\dot{t}^2-\left(1-\frac{2m}{r}\right)^{-1}\dot{r}^2. \tag{6.17}$$

Hence

$$\dot{r}^2=E^2-1+\frac{2m}{r}. \tag{6.18}$$

The rocket can attain a maximum of the parameter r (and then turn back) when $\dot{r}=0$, that is when r takes the value

$$r_M=\frac{2m}{1-E^2}; \tag{6.19}$$

The number r_M is an attained maximum of the parameter r if it is positive and finite, that is if $E<1$. The escape velocity for which r_M is infinite corresponds to $E=1$, hence, by (6.16), to

$$\dot{r}_0^2=\frac{2m}{r_0}. \tag{6.20}$$

The relativistic escape velocity V for the observer at rest is given by the ratio of the space (radial) and time components V^1 and V^0 in the proper frame of this observer of the velocity vector which has components $(\dot{r}_0,\dot{t}_0)$ in the natural frame of the coordinates t,r. The proper frame of the static observer is

$$\theta^0=\left(1-\frac{2m}{r_0}\right)^{\frac{1}{2}}dt,\quad \theta^1=\left(1-\frac{2m}{r_0}\right)^{-\frac{1}{2}}dr, \tag{6.21}$$

therefore

$$V^1=\left(1-\frac{2m}{r_0}\right)^{-\frac{1}{2}}\dot{r}_0,\qquad V^0=\left(1-\frac{2m}{r_0}\right)^{\frac{1}{2}}\dot{t}_0. \tag{6.22}$$

Hence, using (6.20) and (6.18) when $E=1$, we obtain

$$V=:\frac{V^1}{V^0}=\sqrt{\frac{2m}{r_0}}. \tag{6.23}$$

This escape velocity coincides with its Newtonian value. It tends to 1, the light velocity, when r_0 tends to $2m$, in agreement with interpretations that we will give in a forthcoming section.

7 Stability of circular orbits

To study the **stability of circular orbits** one uses the identity

$$g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = -1 \tag{7.1}$$

and Equations (6.4) and (6.6) to obtain

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2m}{r}\right)\left(\frac{\ell^2}{r^2} + 1\right) = \frac{1}{2}E^2. \tag{7.2}$$

This type of differential equation occurs in classical mechanics and governs motion of a particle of unit mass and energy $\frac{1}{2}E^2$ in the potential

$$V(r) \equiv \frac{1}{2}\left(1 - \frac{2m}{r}\right)\left(\frac{\ell^2}{r^2} + 1\right). \tag{7.3}$$

The following second-order differential equation for r can be obtained by differentiating (7.2) with respect to s:

$$\ddot{r} + \frac{dV}{dr} = 0. \tag{7.4}$$

A circular orbit $r = r_0$ is one for which that $\dot{r} = 0$; hence $\ddot{r} = 0$ and $\frac{dV}{dr}(r_0) = 0$. That is, for such an orbit the potential V has a critical point. Computation gives

$$\frac{dV}{dr} \equiv r^{-4}[mr^2 - \ell^2 r + 3m\ell^2]. \tag{7.5}$$

The critical points are therefore given by

$$R_\pm = \frac{\ell^2 \pm \sqrt{\ell^4 - 12\ell^2 m^2}}{2mr^2}. \tag{7.6}$$

For $\ell^2 < 12m^2$ there is no circular orbit. For $\ell^2 > 12m^2$ there are two possible circular orbits: $r_0 = R_+$ and $r_0 = R_-$.

The circular orbit $r = r_0$ is stable if this critical value is a minimum of V, but unstable if it is a maximum. Indeed linearization around r_0 of dV/dr shows that Equation (7.4) leads to oscillations of r around r_0 if $d^2V/dr^2(r_0) > 0$, and exponential growth of r if $d^2V/dr^2(r_0) < 0$. Elementary calculus shows that R_- is a maximum of V and hence the orbit $r_0 = R_-$ is unstable, while R_+ is a minimum and hence $r_0 = R_+$ is stable.

We see from Equation (7.6) that, given m, the smallest possible value of R_+ is $6m$, while the smallest possible value of R_- (obtained when ℓ tends to infinity) is $3m$. We have proved

Theorem 7.1 *In a Schwarzschild spacetime of mass m there is no circular orbit with angular momentum ℓ less than $m\sqrt{12}$. The Schwarzschild coordinates r of circular orbits all satisfy $r > 3m$. The last stable circular orbit has $r > 6m$.*

8 Deflection of light rays

8.1 *Theoretical prediction*

Light rays are null geodesics, so they are bent by the curvature of the spacetime. The differential equations they satisfy are the same as the equations for time-like geodesics, except that the derivative denoted by a dot is now a derivative with respect to the affine parameters on the null geodesic and Equation (7.1) is replaced by

$$g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta = 0. \tag{8.1}$$

That is, using Equations (6.4) and (6.6), we have

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2m}{r}\right)\frac{\ell^2}{r^2} = \frac{1}{2}E^2. \tag{8.2}$$

We deduce from this equation and from Equation (6.4) that on a null geodesic

$$\frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 + \frac{1}{r^2} - \frac{2m}{r^3} = k^2, \tag{8.3}$$

with $k = E^2\ell^{-2}$ a constant. Setting $u \equiv r^{-1}$ gives

$$\left(\frac{du}{d\phi}\right)^2 + u^2 - 2mu^3 = k^2. \tag{8.4}$$

Hence, after differentiation we obtain

$$\frac{d^2u}{d\varphi^2} + u = 3mu^2. \tag{8.5}$$

The term $3mu^2$ in (8.5) is the Einsteinian correction to the following equation:

$$\frac{d^2u}{d\varphi^2} + u = 0, \quad \text{with} \quad u \equiv \frac{1}{r}. \tag{8.6}$$

This gives, in the absence of gravitation, straight lines as light rays, with equations in spherical coordinates taking the form

$$\frac{1}{r} \equiv u_{Str} = \frac{1}{r_0}\cos(\phi - \phi_0), \tag{8.7}$$

where r_0 is the displacement from the centre. If we consider the Einstein correction $3mu^2$ as small we obtain an approximation to Einstein's light rays by solving the differential equation

$$\frac{d^2u}{d\varphi^2} + u = 3mu_{Str}^2. \tag{8.8}$$

Setting $\phi_0 = 0$ for simplicity, we find that the general solution to (8.8) is

$$\frac{1}{r} = \frac{1}{r_0}\cos\phi + \frac{m}{r_0^2}(1 + \sin^2\phi). \tag{8.9}$$

On the straight line (8.7) the coordinate r tends to infinity when ϕ tends to $-\frac{\pi}{2}$ or $\frac{\pi}{2}$. On the curve (8.9), r tends to infinity when ϕ tends to $\pm(\frac{\pi}{2}+\alpha)$ with, by (8.9) and elementary trigonometry, α satisfying

$$-\sin\alpha + \frac{m}{r_0}(1+\cos^2\alpha) = 0. \tag{8.10}$$

That is, if α is small

$$\alpha \cong \frac{2}{r_0} \tag{8.11}$$

The total deflection of a light ray is therefore estimated to be

$$\delta \cong \frac{4m}{r_0}. \tag{8.12}$$

For light rays grazing the solar surface this number is about 1.75 seconds of arc.

8.2 *Results of observation*

It was, until recently, necessary to wait for an eclipse to be able to observe the apparent displacement of the stars due to the deflection by the Sun of light coming from them. As early as 1919 an expedition was organized by Eddington to measure the bending of light by the Sun. The observed deflection was in reasonable agreement with the prediction. The observation of light deflection has not to date been very precise, in spite of a number of subsequent attempts (including an expedition to Mauritania in 1973). The experimental error is still estimated to be more than 10 per cent.

More precise results have been obtained with the use of radio waves, since it is not necessary in this case to wait for an eclipse. Measurements using several radio telescopes at intercontinental distance (Very Long Base Interferometry) give a precision of about 10^{-4}. They strongly support General Relativity, in particular in contrast to the Brans–Dickie scalar-tensor theory.

8.3 *Fermat's principle and light travel parameter time*

We have said that light rays follow null geodesics of spacetime. The following property is a generalization to static spacetimes of a classical theorem of Fermat in Newton's space E^3.

Theorem 8.1 *In a static spacetime with metric*

$$ds^2 = -g_{00}dt^2 + g_{ij}dx^i dx^j, \tag{8.13}$$

the projections on space of light rays are geodesics of the Riemannian metric

$$d\sigma^2 = \frac{g_{ij}dx^i dx^j}{g_{00}}; \tag{8.14}$$

i.e. they are relative minima of the integral

$$\ell = \int_{t_1}^{t_2} \sqrt{\frac{g_{ij}dx^i dx^j}{g_{00}}}. \tag{8.15}$$

On a null curve ℓ is the parameter time duration.

The proof is a straightforward computation, using the geodesic equations in ds and $d\sigma$. We leave it as an exercise.

In addition we note that Fermat's principle can be generalized to arbitrary spacetimes as follows[5]:

Theorem 8.2 *The "proper time of arrival" of null curves to a given timelike curve admits a critical point (is a relative minimum) at a light ray.*

9 Red shift and time delay

The time delay and redshift effects are not due to a change of the velocity of light in a gravitational field, since in General as in Special Relativity the speed of light is a universal constant. Rather the effects are caused by dependence on the observer of the measure of the proper time (see Chapter 3).

The red shift

We have already considered the gravitational red shift in Chapter 3, in any static spacetime. Its measurement in a Schwarzschild spacetime is one of the classical tests of Einstein's equations.

Suppose for simplicity that the emitting atom and the observer are both at rest, with the same angular coordinates and radial Schwarzschild coordinates, respectively r_A and r_O. The emitted period T_A and the observed periods T_O are then linked by the relation (see equation III 4.5)

$$T_O = \left(1 - \frac{2m}{r_O}\right)\left(1 - \frac{2m}{r_A}\right)^{-1} T_A. \tag{9.1}$$

Hence $T_O > T_A$ if $r_O > r_A$, so a red shift (smaller frequency ν) of spectral lines is then observed. If m/r_0 is small then

$$\frac{T_A}{T_O} \equiv \frac{\nu_O}{\nu_A} \cong 1 - 2m\left(\frac{1}{r_A} - \frac{1}{r_O}\right). \tag{9.2}$$

Time delay

To state the elements of the theoretical time delay predictions we treat the case of a rocket sent from the Earth in a radial direction with a velocity less than the escape velocity and moving then in free fall, i.e. following a timelike geodesic of the Schwarzschild spacetime representing the Earth's gravitational field.

[5] Ferrarese, G. (2004) *Riferimenti Generalizzati in Relatività e Applicazioni* Pitagora editrice, Bologna.

The parameter time t_M it takes the rocket to fall back on Earth is twice the parameter time it takes to attain its maximum r_M value (see Section 6.3); we calculate the proper time using the formula

$$s_A = \int_0^{t_M} \frac{ds}{dt} dt = \int_0^{t_M} E^{-1}\left(1 - \frac{2m}{r}\right) dt = \int_0^{t_M} \frac{1 - \frac{2m}{r}}{1 - \frac{2m}{r_0}} \frac{ds}{dt}(0) dt \tag{9.3}$$

with

$$\left(\frac{ds}{dt}(0)\right)^2 = \left(1 - \frac{2m}{r_0}\right) - \left(1 - \frac{2m}{r_0}\right)^{-1} \left(\frac{dr}{dt}(0)\right)^2. \tag{9.4}$$

We set $dr/dt(0) = v$ and we obtain for s_A, when m/r_0 is small the approximate expression

$$s_A \tilde{=} \int_0^{t_M} \left(1 - \frac{2m}{r} + \frac{m}{r_0} + \frac{mv^2}{r_0}\right) dt. \tag{9.5}$$

The proper time observed on Earth at the fixed r_0 between the origin and the impact point is

$$s_O = \int_0^{t_M} \sqrt{1 - \frac{2m}{r_0}} dt \cong \int_0^{t_M} \left(1 - \frac{m}{r_0}\right) dt. \tag{9.6}$$

Therefore a standard clock carried by the rocket shows a delay over the same standard clock of the observer, approximately

$$s_A - s_O \cong \int_0^{t_M} \left(-\frac{2m}{r} + 2\frac{m}{r_0} + \frac{mv^2}{r_0}\right) dt > 0, \tag{9.7}$$

since on the trajectory $r > r_0$. Experiments made with caesium clocks have confirmed this time delay and its estimates with great precision.

Remark 9.1 *On a Lorentzian manifold timelike geodesics are, as in Minkowski spacetime, local maxima of the length of timelike curves joining two points. In the given example it is the rocket which follows a geodesic, while in the twin paradox (see Chapter 2) it is the travelling one which is not in free fall, as he has to use a motor to come back.*

10 Spherically symmetric interior solutions

A relativistic model for a spherically symmetric isolated star is a spherically symmetric spacetime $(V \equiv R^3 \times R, g)$ such that g satisfies the Einstein equations on V

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = T_{\alpha\beta}$$

where $T_{\alpha\beta}$, the source stress energy tensor, is zero outside of the star and depends inside on the physical constitution of the star. This is not very well known, and

difficult to model in a single formula. For this reason there is no physically reliable exact solution for the considered problem. However, important qualitative features can be obtained from general considerations.

Remark 10.1 *As in Newtonian theory, a spherically symmetric Einsteinian gravitational field vanishes in the interior of a hollow sphere, namely a spherically symmetric vacuum solution defined for $r < a$ is necessarily flat in this domain. Indeed, the construction of Section 3 shows that this solution is a Schwarzschild metric (3.1), with m some arbitrary number. The only solution of this form which is continuous at $r = 0$ is the flat solution with $m = 0$.*

10.1 *Static solutions. Upper limit on mass*

We look for equilibrium configurations, i.e. static spacetimes. We use Schwarzschild coordinates. For $r > r_{star}$, with r_{star} the radial Schwarzschild coordinate of the star boundary, we have vacuum; hence the solution is Schwarzschild. For $r < r_{star}$ we again look for a metric of the form

$$g = -e^{\nu}dt^2 + e^{\lambda}dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2),$$

but now we solve the equations with a non-vanishing stress energy tensor which we take, in the absence of better modelling, to represent a perfect fluid:

$$T_{\alpha\beta} \equiv (\mu + p)u_\alpha u_\beta + pg_{\alpha\beta}.$$

Then the Einstein equations are

$$R_{\alpha\beta} = \rho_{\alpha\beta} \equiv (\mu + p)u_\alpha u_\beta + \frac{1}{2}(\mu - p)g_{\alpha\beta}. \tag{10.1}$$

For a static solution, we have the fluid unit velocity, $u^i = 0$, $u^0 = e^{-\nu/2}$, tangent to the time lines; then

$$T_{00} \equiv \mu e^{\nu}, \quad T_{0i} = 0, \quad T_{ij} \equiv pg_{ij}. \tag{10.2}$$

The scalar functions μ and p, like λ and ν, depend only on r.

Using the identities

$$R_{22} + \frac{1}{\sin^2\theta}R_{33} \equiv 2\left\{-e^{-\lambda}\left(1 + \frac{r}{2}(\nu' - \lambda')\right) + 1\right\}$$

and

$$r(e^{\lambda-\nu}R_{00} + R_{11}) \equiv \nu' + \lambda',$$

we see that the scalar curvature of the metric g is

$$R \equiv -e^{-\nu}R_{00} + e^{-\lambda}R_{11} + r^{-2}(R_{22} + sin^{-2}\theta R_{33}). \tag{10.3}$$

$$\equiv -e^{-\lambda}\left(\nu'' + \frac{\nu'^2}{2} - \frac{\nu'\lambda'}{2} + \frac{\nu' - \lambda'}{r}\right) \tag{10.4}$$

We then find that the equation (Hamiltonian constraint: by the general theory – see Chapter 6 – S_0^0 does not contain the lapse, here e^ν)

$$S_{00} \equiv R_{00} - \frac{1}{2} g_{00} R = T_{00} \equiv \mu e^\nu$$

reduces therefore to the following differential equation for λ :

$$e^{-\lambda}(r\lambda' - 1) + 1 = 2r^2\mu, \tag{10.5}$$

which we can rewrite

$$\frac{d}{dr}[r - re^{-\lambda}] = 2r^2\mu. \tag{10.6}$$

Smoothness of the metric in a neighbourhood of the origin $r = 0$ imposes the vanishing of the integration constant; hence

$$e^{-\lambda} = 1 - \frac{2M(r)}{r} \quad \text{with} \quad M(r) \equiv \int_0^r \rho^2 \mu(\rho) d\rho. \tag{10.7}$$

Remark 10.2 *For a smooth density μ we require that e^λ tends to 1 when r tends to zero, as needed for the smoothness of a metric in the neighbourhood of the origin of spherical "coordinates".*

Remark 10.3 *The solution can define a Lorentz metric in a domain $r \leq a$ only if*

$$M(a) \equiv \int_0^a \rho^2 \mu(\rho) d\rho \leq \frac{a}{2}. \tag{10.8}$$

The number $M(a)$ is not the total proper mass of a star of standard radius a. Indeed, this proper mass must be computed with the proper (geometric) volume element, that is[6]

$$M_a := \int_0^a \rho^2 \mu(\rho) \left(1 - \frac{2M(\rho)}{\rho}\right)^{-\frac{1}{2}} d\rho, \tag{10.9}$$

We see that the proper mass M_a is greater than $M(a)$; the difference represents the gravitational binding energy of the star in equilibrium.

Having computed λ, we use the equation

$$\nu' + \lambda' \equiv r(e^{\lambda-\nu} R_{00} + R_{11}) = r(e^{\lambda-\nu}\rho_{00} + \rho_{11}) \equiv re^\lambda(\mu + p) \tag{10.10}$$

and the expression of λ' deduced from (10.3) and $e^\lambda = \frac{r}{r-2M}$. We find that:

$$\nu' = (e^\lambda - 1)r^{-1} + re^\lambda(p - \mu) = \frac{2M(r) + r^3(p - \mu)}{r[r - 2M(r)]}. \tag{10.11}$$

[6] The factor 4π which should appear in the integral over angular variables has been incorporated, in our convention for geometrical units, in the proper density μ.

On the other hand, using the fluid equation (cf. Chapter 3)

$$(\mu + p)u^\alpha \nabla_\alpha u_1 + \partial_1 p \equiv (\mu + p)u^0 u_0 \Gamma^1_{01} + \partial_1 p = 0,$$

i.e.

$$p' = -\frac{1}{2}(p + \mu)\nu',$$

we find that:

$$p' = -\frac{1}{2}(p + \mu)\frac{2M(r) + r^3(p - \mu)}{r[r - 2M(r)]}. \tag{10.12}$$

This is known as the **Tolman–Oppenheimer–Volkov** equation of relativistic hydrostatic equilibrium. It reduces to the Newton equation for hydrostatic equilibrium when p is negligible with respect to μ and $M(r)$ negligible with respect to r. The equation could be integrated if we knew an equation of state inside the star, but this is not physically realistic due to the complexity of the phenomena which take place. However, one can deduce some conclusions from this formula.

When the density μ is a constant, μ_0, the integration of (10.7) gives

$$M(r) = \mu_0 \int_0^r \rho^2 d\rho = \frac{1}{3}\mu_0 r^3. \tag{10.13}$$

Equation (10.12) can then be integrated exactly, to give[7] for the pressure at the centre of the star

$$p(0) = \mu_0 \frac{1 - (1 - 2a^{-1}M(a))^{\frac{1}{2}}}{3[1 - 2a^{-1}M(a)]^{\frac{1}{2}} - 1}. \tag{10.14}$$

This pressure becomes infinite if

$$M(a) = \frac{4a}{9}, \tag{10.15}$$

and negative if $M(a) > \frac{4a}{9}$. The classical conclusion is that stars with $M(a) \geq \frac{4a}{9}$ cannot exist.

The same result holds if one supposes only that μ is a non-increasing function of r. We enunciate a theorem.

Theorem 10.4 *There exist no equilibrium configuration of spherically symmetric stars filled with a perfect fluid with μ a non-increasing function of r and such that*

$$M(a) \geq \frac{4a}{9}. \tag{10.16}$$

where a is the standard radius of the star and $M(a)$ the integral of its density.

[7] Integration first performed by Schwarzschild in 1916. For details, and discussion of the case of a non-constant density, see Wald (1984), *General Relativity*, University of Chicago Press, p. 129.

10.2 *Matching with an exterior solution*

An Einsteinian model for the exterior and the interior of a spherically symmetric star is a manifold $R^3 \times R$ with a Lorentzian metric g which satisfies the Einstein equations on the whole manifold, and induces an interior Schwarzschild metric in $B \times R$, with B a ball of R^3, and an (exterior) Schwarzschild metric in the complementary domain. Since the coordinates have a geometrical meaning, they should extend across the star boundary. The full metric reads therefore

$$g = -e^{\nu}dt^2 + e^{\lambda}dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2),$$

with e^{λ} and e^{ν} equal to the coefficients of the interior metric g_{Star} for $r < r_{Star}$, and to those of an exterior Schwarzschild metric

$$g_m = -\left(1 - \frac{2m}{r}\right)dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2)$$

for $r > r_{Star}$. The stress energy tensor is zero for $r > r_{Star}$ and given by (10.2) for $r < r_{Star}$, and is therefore continuous by pieces. The same must be true of the Ricci tensor of the full metric. The expressions given in Section 3 for the components of the Ricci tensor show that this requirement on the Ricci tensor implies the continuity of λ across[8] $r = r_{Star}$; that is,

$$m = M(r_{Star}) \equiv \int_0^{r_{Star}} r^2\mu(r)dr.$$

As we remarked before the m so obtained is greater than the total proper mass of the star as computed with the proper volume element. This reflects the fact that the gravitational field itself contributes to the "mass". We have proved the following theorem.

Theorem 10.5 *There are spherically symmetric metrics g continuous on $R^3 \times R$ which satisfy the Einstein equations with source a perfect fluid of pressure p and density μ contained in the ball $r \leq r_{Star}$.*

Exercise. Write the relation between the Schwarzschild mass m *in the exterior of the star; $r > r_{Star}$, and the interior quantities.*

10.3 *Non-static solutions*

The Birkhoff theorem does not apply to interior solutions. There exist time-dependent, spherically symmetric solutions of the Einstein equations with sources in a domain $r < r_{Star}$, and the radius r_{Star} of the star may be a function of t. The full solution is the considered interior solution g_{int} for $r \leq r_{Star}$ and a Schwarzschild exterior solution with mass m for $r \geq r_{Star}$ as long as $r_{Star} > m$. The number m, called the Schwarzschild radius of the star, is a constant for an

[8] If these quantities were not continuous, their derivation would introduce a measure with support $r = r_{Star}$; see Chapter 3.

exterior solution, linked with the energy content of the star which is considered as an isolated object. When the star contracts so much that r_{Star} becomes smaller than m, it is no longer observable from the region $r > m$ of the spacetime (see next section).

It is physically clear that there can be no static interior solution with zero pressure, because there is nothing in that case to resist the gravitational attraction. This can be checked mathematically by solving Equation (10.12) with $p = p' = 0$, which reads

$$\mu(r) = 2r^{-3}M(r) \equiv 2r^{-3}\int_0^r \rho^2\mu(\rho)d\rho, \tag{10.17}$$

that is, setting $y(r) = \int_0^r \rho^2\mu(\rho)d\rho$, the differential equation

$$y' = \frac{2}{r}y, \quad \text{hence} \quad y = Ce^{r^2} \tag{10.18}$$

with C a constant. Then

$$\mu = Cr^{-2}e^{r^2}. \tag{10.19}$$

always infinite for $r = 0$.

We will return in Section 12 to the **gravitational collapse**.

11 The Schwarzschild black hole

11.1 *The event horizon*

The Schwarzschild spacetime is defined only for $r > 2m$. The boundary of the manifold $r > 2m$ in $R^3 \times R$ is the submanifold $r = 2m$ of $R^3 \times R$, diffeomorphic to a product[9] $S^2 \times R$. This submanifold is called the **event horizon**, or simply the horizon, for reasons we will justify in the next section.

We study the approach to the horizon of timelike and null radial geodesics.

Radial curves (i.e. $\theta =$ constant and $\varphi = constant$; hence $\ell = 0$) are geodesics on which the element of length is given by

$$ds^2 = (1 - 2mr^{-1})dt^2 - (1 - 2mr^{-1})^{-1}dr^2.$$

Hence the radial light rays satisfy (for $r > 2m$)

$$\frac{dr}{dt} = \pm(1 - 2mr^{-1})$$

which integrates to give the ingoing and outgoing light rays, corresponding respectively to the minus and the plus sign, issued from a point (r_0, t_0):

$$t - t_0 = \pm(r - r_0 + 2m\,\log\left[\left(\frac{r}{2m} - 1\right)\right]\left[\left(\frac{r_0}{2m} - 1\right)\right]^{-1}$$

[9] Recall that r is a polar coordinate on R^3 and S^2 is a topological sphere.

The ingoing rays always tend asymptotically to the line $r = 2m$ in the (t, r) plane when t tends to infinity. In other words t becomes infinite as $r > 2m$ decreases to $2m$ along this light ray, and a static observer (i.e. moving on a line where t alone is varying) will never receive a signal saying that the light ray has attained the boundary $r = 2m$. Hence no future directed causal curve attains, in a finite standard parameter time, a point where $r = 2m$.

11.2 *The Eddington–Finkelstein extension*

The Schwarzschild metric in standard coordinates, with $m > 0$, ceases to be a smooth Lorentzian metric for $r = 2m$; for such a value of r the coefficient g_{00} vanishes while g_{11} becomes infinite. For $0 < r < 2m$ it is again a smooth Lorentzian metric but t is a space coordinate while r is a time coordinate. Hence the metric cannot be said to be either spherically symmetric or static for $r < 2m$.

It is immediate to see that the determinant $-r^4 sin^2\theta$ of g_{Schw} is regular (smooth and non-vanishing) for $r = 2m$. In addition a straightforward computation shows that the coordinate-independent scalar $R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta}$, called **Kretschmann** scalar, is finite at $r = 2m$; in fact it is equal to $48m^2/r^6$. This property led to the belief that the Schwarzschild spacetime $(S^2 \times \{r > 2m\}) \times R$ is extendible, in the sense that it can be immersed in a larger spacetime whose manifold is not covered by the Schwarzschild standard coordinates with $r > 2m$.

One considers (Eddington, 1926; Finkelstein, 1958) the change of coordinates obtained by replacing the canonical Schwarzschild time by the "retarded time" v given when $r > 2m$ by (cf. equation of radial null geodesics)

$$v = t + r + 2m \ln\left(\frac{r}{2m} - 1\right). \tag{11.1}$$

The Schwarzschild metric reads in this new coordinate

$$-\left(1 - \frac{2m}{r}\right) dv^2 + 2drdv + r^2(sin^2\theta d\varphi^2 + d\theta^2). \tag{11.2}$$

The above metric is defined on the manifold $S^2 \times (r > 0) \times R$ and it is a regular Lorentzian metric: its coefficients are smooth and the term $drdv$ ensures its non-degeneracy for $r = 2m$. The corresponding Eddington–Finkelstein spacetime is an extension of the Schwarzschild spacetime, because the open subset $r > 2m$ of that spacetime is isometric to the Schwarzschild spacetime. The Eddington–Finkelstein spacetime is stationary but not static: it is not invariant under time reversal. It is time oriented by decreasing v.

The submanifold $r = 2m$ is null (isotropic), since we find that $g^{rr} = 0$ for $r = 2m$.

One family of radial (i.e. $\theta = \text{constant}$, $\phi = \text{constant}$) light rays is represented by straight lines $v = \text{constant}$; the other family is given by

$$v = 2r + 4m \log|r - 2m| + constant. \tag{11.3}$$

Exercise. *Show that on the domain $r < 2m$ of the Eddington–Finkelstein spacetime the metric takes the Schwarzschild form by the change of coordinates $t = v - r - 2m \log(1 - \frac{r}{2m})$.*

Theorem 11.1 *On a timelike or null line issued from a point with $r < 2m$ the variable r is always decreasing and tends to zero in a finite proper time. Hence any observer crossing the Schwarzschild radius $r = 2m$ attains the singularity $r = 0$ in a finite proper time.*

Proof The proof is conveniently done in terms of the t, r coordinates, notwithstanding the fact that t is now a space coordinate and r a time one.

1. Radial geodesics. We have found that on such geodesics

$$\dot{r}^2 = E^2 - (1 - 2mr^{-1}). \tag{11.4}$$

Hence the point $r = 0$ is attained from $r = r_0 < 2m$ after the finite proper time

$$\int_0^{r_0} \frac{r^{\frac{1}{2}} dr}{\sqrt{2m - (1 - E^2)r}}. \tag{11.5}$$

2. Non-radial geodesics and non geodesic motions lead to a smaller lapse of proper time between points of coordinates $r_0 < 2m$ and 0. □

Remark 11.2 *The amount of proper time for an astronaut entering the black hole to reach the singularity is indeed very short. It is of the order of the Schwarzschild radius when the light velocity is taken equal to 1. For a black hole with mass of the order of the solar mass it gives a time of the order of $10^{-5}s$.*

11.3 *Eddington–Finkelstein white hole*

By time reversal one obtains manifestly another extension of the Schwarzschild space time. The manifold is again $S^2 \times (r > 0) \times R$ and the metric is

$$-\left(1 - \frac{2m}{r}\right) dv^2 - 2drdv + r^2(sin^2\theta d\varphi^2 + d\theta^2).$$

The extension to $r < 2m$ appears now to observers in the Schwarzschild spacetime as a white hole: nothing can penetrate into it, but every past inextendible light ray or time line in the Schwarzschild spacetime emanates from this white hole.

11.4 *Kruskal complete spacetime*

It is possible to embed the Schwarzschild spacetime and both of its extensions in a larger spacetime containing an additional copy of the Schwarzschild space time for which the previous black hole extension now plays the role of a white hole and vice versa. The support of the obtained spacetime is the manifold $S^2 \times R^2$.

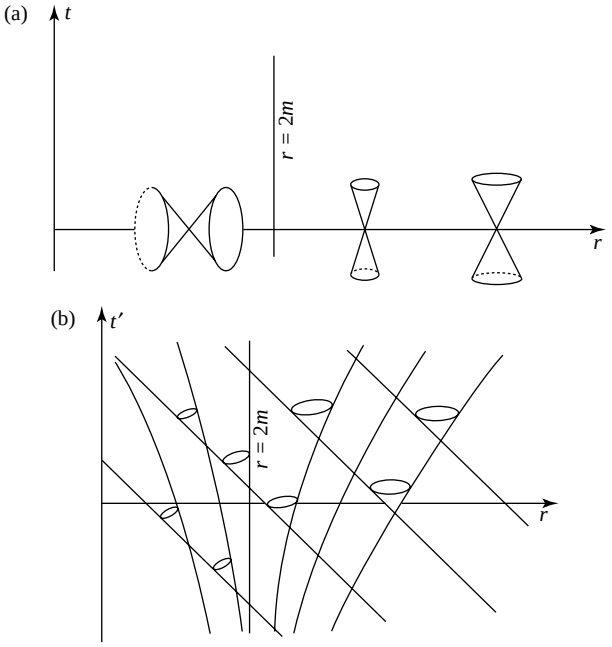


Figure 4.1 (a) The orientation of the light cones; (b) null geodesics and light cones in Eddington–Finkelstein extension.

11.4.1 *Kruskal–Szekeres coordinates*

The metric of the Kruskal spacetime reads in the Kruskal coordinates

$$\frac{32m^3}{r}e^{-r/2m}[dz^2 - dw^2] + r^2(d\theta^2 + \sin^2\theta d\phi^2),$$

where θ, ϕ are coordinates on S^2 while z, w are coordinates on the open set diffeomorphic to R^2 defined by

$$z^2 - w^2 > -1,$$

and r is the function of z and w defined by

$$z^2 - w^2 = \frac{1}{2m}(r - 2m)e^{r/2m}.$$

In Kruskal coordinates the radial light rays are represented by straight lines.

The Kruskal spacetime has two asymptotically flat regions, each isometric to the Schwarzschild spacetime. A section through the Kruskal spacetime connecting these two regions, for instance $w = 0$, is called the Schwarzschild throat. No signal can travel from one of the asymptotically flat regions to the other.

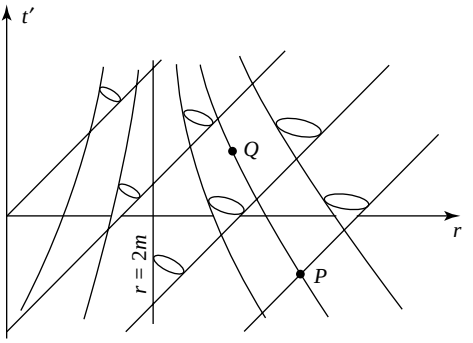


Figure 4.2 Radial null geodesics and light cones of the white hole.

Exercise. Obtain the portion of the Kruskal spacetime isometric to the Schwarzschild spacetime by the change of coordinates:

$$z = e^{v/4m} + e^{-u/4m}, \qquad w = e^{v/4m} - e^{-u/m}$$

with

$$u = t - r - 2m\ log\left(\frac{r}{2m} - 1\right).$$

Obtain other portions by analogous changes of coordinates.

11.5 *Observations*

Laplace and Mitchell had already foreseen that light can be trapped by a massive body, so that it becomes black to observers and is perceptible only through its gravitational field. However, the apparent "Schwarzschild singularity" $r = 2m$ is a phenomenon which has no analogue in classical mechanics, and the completed spacetimes are indeed very strange. It was long believed that these extensions had no physical reality, i.e. that matter cannot be compressed so that its Schwarzschild radius is less than its mass. It is mainly through the vision of J. A. Wheeler that the reality of black holes was seriously considered. They cannot be seen directly, but astronomical observations, in particular perturbations of motions of various stars, have led to a large consensus on the existence of many black holes in the Universe, even at the centre of our own galaxy. We will return to more general black holes in another chapter.

12 Spherically symmetric gravitational collapse

It is believed that when a star has exhausted all thermonuclear sources of energy, it will collapse under its own gravitational field. Oppenheimer and Snyder[10] have given a rigorous study of the process for a spherically symmetric dust cloud by

[10] Oppenheimer, J. R. and Snyder, H. (1939) *Phys. Rev.* **56**, 455.

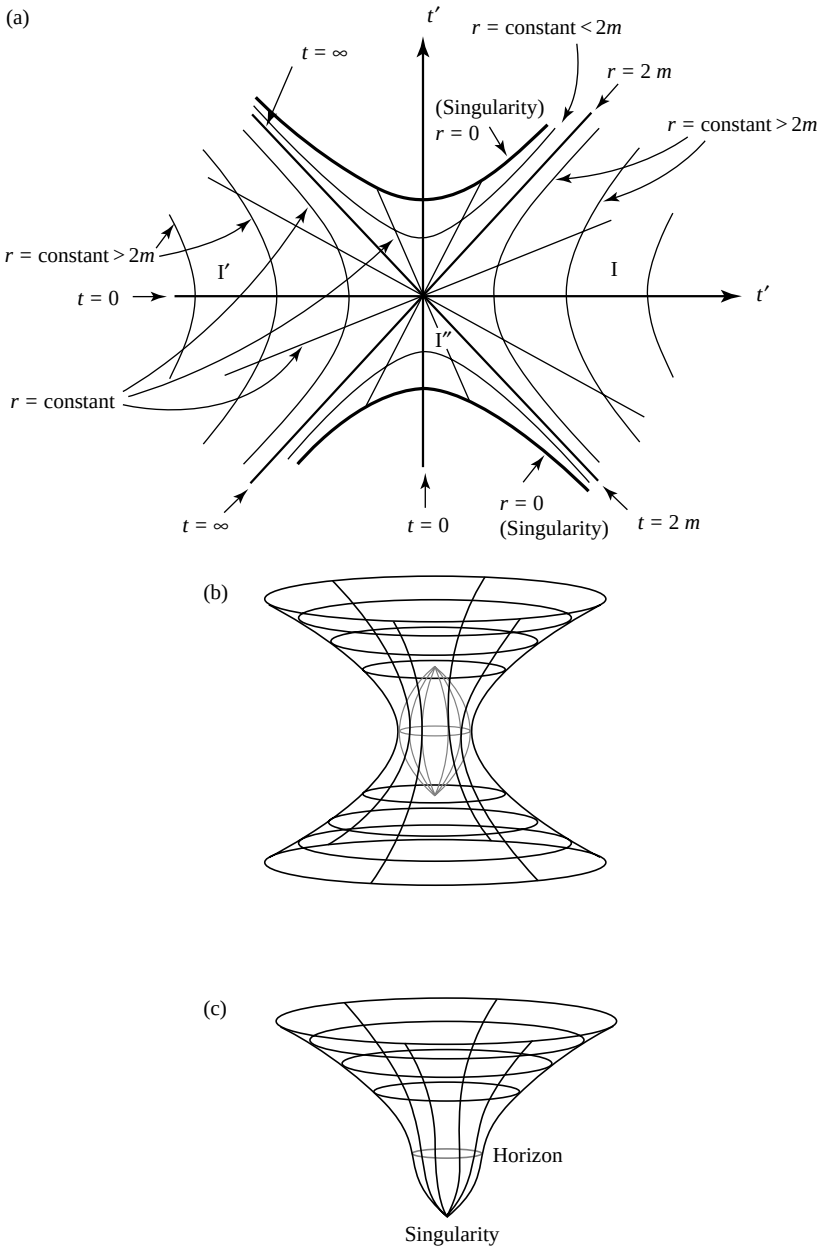


Figure 4.3 Kruskal spacetime and Schwarzchild black hole.

using a special solution given by Tolman[11], in the restricted case of a spatially constant matter density starting from rest. This study was taken anew by Gu Chao Hao[12], in a study of the Tolman solution, but not restricted to spatially constant matter. Gu Chao Hao gave an analysis of shell crossing singularities as well as the essential singularity at the centre of symmetry. He gave a detailed mathematical discussion of the formation of horizons and of what is now called the nakedness of these singularities. Hu Hesheng[13] constructed the complete solutions of the fully general spherically symmetric solution for dust, not restricted to the Tolman class, and discussed their properties. The work done by Gu Chao Hao and Hu Hesheng for non-homogeneous dust clouds was done again, unaware of these papers published in China, by Müller zum Hagen, Seifert, and Yodzis[14] (Tolman class, shell crossing singularities), Christodoulou[15] (starting from rest, central singularities) and Richard Newman[16] (general solution). These western papers were motivated by the cosmic censorship conjecture, which is violated both at "shell crossing" singularities and at the centre, as was already shown in Gu's paper.

12.1 *Tolman metric*

In the case of dust the flow lines are the timelike geodesic trajectories of the particles; we take these geodesics as timelines (comoving coordinates). A spherically symmetric metric can then be written as

$$-dt^2 + e^{2\omega}dr^2 + R^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{12.1}$$

where ω and R are functions of r and t. The flow lines of the dust are the curves on which the coordinates r, θ, ϕ are constant; a so-called dust shell is labelled by its coordinate r. The metric is regular as long as $e^{2\omega}$ and R are smooth positive functions. R^2 is – up to multiplication by 4π – the area of the spherical dust shell of parameter r at time parameter t. One fixes the parameter r by choosing it to be such that $r = R(r, 0)$.

The stress energy tensor T of the dust reduces to

$$T_{\alpha\beta} = \mu u_\alpha u_\beta, \tag{12.2}$$

with $u^\alpha = \delta^\alpha_0$, since the time lines are the matter flow lines. The only non-zero component of the stress energy tensor is

$$T_{00} = \mu, \tag{12.3}$$

[11] Tolman, R. (1934) *Proc. Natl. Acad. Sci. USA*, **20**, 169–76, Bondi, H. (1947) *Mon. Nat. R. Astron. Soc.*, **107**, 410–25.
[12] Gu Chao Hao (1973) *J. Fudan University*, **1**, 73–8.
[13] Hu Hesheng (1974) *J. Fudan University*, **2**, 92–8.
[14] Müller zum Hagen, H., Seifert, H. J., and Yodzis, P. (1973) *Comm. Math. Phys.*, **34**, 135–48).
[15] Christodoulou, D. (1984) *Commun. Math. Phys.*, **93**, 171–95.
[16] Newman, R. (1986) *Class. Quantum Grav.*, **3**, 527–39.

with

$$\mu = \mu(r,t) \quad for \quad r \leq a \quad \text{and} \quad \mu(r,t) = 0 \quad for \quad r > a, \tag{12.4}$$

where a is the r parameter of the outermost dust shell[17] at time $t = 0$.

One denotes respectively by primes and dots differentiation with respect to r and t. We compute the general solution in the chosen notation. The equation $R_{10} = 0$ is equivalent to the second-order equation

$$\dot{R}' - \dot{\omega} R' = 0, \tag{12.5}$$

which integrates to the first-order equation

$$R' e^{-\omega} = f(r), \tag{12.6}$$

with $f(r)$ an arbitrary function of r.

The conservation equations reduce to

$$\nabla_\alpha T^{\alpha 0} \equiv \dot{\mu} + \Gamma^\alpha_{\alpha 0} \mu \equiv \dot{\mu} + \mu(\dot{\omega} + 2R^{-1}\dot{R}) = 0; \tag{12.7}$$

Hence, with ϕ an arbitrary function of r, one obtains

$$\mu(t,r) = \frac{e^{-\omega}}{R^2}\phi(r). \tag{12.8}$$

Using (12.6), this equation becomes

$$\mu(t,r) = \frac{r^2 \mu_0}{R^2 R'}. \tag{12.9}$$

Here μ_0 is an arbitrary function of r which we identify with the initial density, because we have chosen r such that $R(0,r) = r$, hence $R'(0,r) = 1$. It has been shown by R. Newman that the Einstein equations admit then the first integral

$$\frac{1}{2}\dot{R}^2 - \frac{M(r)}{R} = \frac{1}{2}[f^2(r) - 1], \tag{12.10}$$

with

$$M(r) = \int_0^r f(\rho)\mu(t,\rho)R^2(t,\rho)e^{\omega} d\rho. \tag{12.11}$$

The function $M(r)$ is the integral over the volume occupied by the dust shells $0 \leq \rho \leq r$ of the density weighted by the factor $f(\rho)$[18]. It is independent of t, as can be verified by using equations (12.6) and (12.9), which show that:

$$M(r) = \int_0^r \mu_0(\rho)\rho^2 d\rho. \tag{12.12}$$

[17] Which may not remain the outermost shell if there is shell crossing as time evolves.

[18] Remark as in a previous section that a factor 4π has been included in our definition of densities; the integral giving $M(r)$ is taken over a 3-sphere.

The explicit formulae for $R(r,t)$ and $\omega(r,t)$ depend on the sign of $f^2(r)-1$. They have been computed and studied in the general case by Hu Hesheng (1974) and Newman (1986) (see above), to whom we refer the reader. The case where the dust cloud starts from rest, i.e.

$$\dot{R}(0,r) = 2\frac{M(r)}{r} + f^2(r) - 1 = 0, \tag{12.13}$$

has been studied by Christodoulou, using parametric equations for R and t.

Here, following Gu Chao Hao (1973), we study the case $f^2(r)=1$, in which exact integration permits clear and comparatively short discussions. In the case $f^2(r)=1$ Equation (12.10) reads

$$R^{\frac{1}{2}}\dot{R} = \pm\sqrt{2M(r)}. \tag{12.14}$$

We suppose that the star starts contracting; then the minus sign must be chosen when t increases and the equation integrates to

$$R(r,t)^{\frac{3}{2}} = \Phi(r) - \frac{3}{2}\sqrt{2M(r)}t, \tag{12.15}$$

with Φ an arbitrary function. Since we have normalized the radial parameter r by the condition $R(r,0)=r$, it holds that $\Phi(r)=r^{\frac{3}{2}}$. We use the notation

$$\frac{3}{2}\sqrt{2M(r)} := h^{\frac{1}{2}}(r);$$

Then (12.15) reads

$$R(r,t) = \{r^{\frac{3}{2}} - h^{\frac{1}{2}}(r)t\}^{\frac{2}{3}}. \tag{12.16}$$

Using Equation (12.6) we find then

$$e^{\omega} = R' = \{r^{\frac{3}{2}} - h^{\frac{1}{2}}(r)t\}^{-\frac{1}{3}}\{r^{\frac{1}{2}} - \frac{1}{3}h^{-\frac{1}{2}}h'(r)t\}.$$

Here, as a consequence of the definition (12.15), we have

$$h'(r) = \frac{9}{2}r^2\mu_0(r).$$

On the other hand, the equation for μ gives

$$\mu(t,r) = \frac{r^2\mu_0}{\{r^{\frac{3}{2}} - h^{\frac{1}{2}}(r)t\}\{r^{\frac{1}{2}} - \frac{3}{2}h^{-\frac{1}{2}}h'(r)t\}}. \tag{12.17}$$

We make a change of coordinates: instead of choosing as radial parameter the number characterizing a dust shell we take the area R of a dust shell at time t. Since we have chosen $f(r)=1$, i.e. $R'=e^{\omega}$, it holds that:

$$dR = e^{\omega}dr + \dot{R}dt = e^{\omega}dr - R^{-\frac{1}{2}}\sqrt{2M(r)}dt. \tag{12.18}$$

The change of coordinates from r, t to R, t is admissible if R' does not vanish. In the new coordinates the spacetime metric becomes, for $0 \leq r \leq a$,

$$-\left(1 - \frac{\sqrt{2M(r)}}{R}\right) dt^2 + 2R^{-\frac{1}{2}}\sqrt{2M(r)}dRdt + dR^2 + R^2(d\theta^2 + sin^2\theta d\phi^2). \tag{12.19}$$

We see that the metric of the space sections $t =$ constant reduces to the Euclidean metric in polar coordinates. If the space manifold is, for every t, homeomorphic to a ball of three-dimensional Euclidean space and R is an admissible polar coordinate, then $R(t, r)$ is also the distance at time t between the centre and the dust shell with parameter r.

12.2 *Monotonically decreasing density*

12.2.1 *Collapse of the dust shells*

In this section we suppose that μ_0 is a monotonically decreasing function of the parameter r as r increases from 0, the centre of the star, to a, the value of r at the surface boundary of this star at time $t = 0$.

To study the possible collapse we study the evolution in proper time t of the function $R(t, r)$ for a given dust shell, i.e. for a given value of r. We see from (12.16) that the shell collapses at the centre of symmetry at the time $t_1(r)$ where $R(t_1, r) = 0$; that is

$$t_1(r) = h^{-\frac{1}{2}}(r)r^{\frac{3}{2}}.$$

It follows that

$$\frac{dt_1(r)}{dr} = \frac{1}{2}r^{\frac{1}{2}}h^{-\frac{3}{2}}\{3h(r) - rh'(r)\} = \frac{9}{4}r^{\frac{1}{2}}h^{-\frac{3}{2}}\{3M(r) - rM'(r)\}. \tag{12.20}$$

If μ_0 is monotonically decreasing, using the definition 12.11 of $M(r)$, we find that

$$3M(r) - rM'(r) \equiv 3\int_0^r \mu_0(\rho)\rho^2 d\rho - \mu_0(r)r^3 \geq 0. \tag{12.21}$$

Hence the shells with increasing parameter r arrive successively at the centre, and there is no shell crossing[19].

Remark. If the density is uniform (Oppenheimer–Snyder case), the dust shells all arrive at the same time at the centre.

The metric (12.19) is a regular Lorentzian metric if the linear form $dR - \dot{R}dt$ does not vanish, i.e. if $R' = e^{\omega} > 0$. We deduce from the expression (12.16) of R that

$$\frac{3}{2}R^{\frac{3}{2}}R' = \left\{\frac{3}{2}r^{\frac{1}{2}} - \frac{1}{2}h^{-\frac{1}{2}}(r)h'(r)t\right\} = 0. \tag{12.22}$$

[19] Shell crossing, leading to non-central singularities, exists when the density is not monotonically decreasing (see Gu Chao Hao and other quoted references).

Hence $R'(t,r)$ vanishes[20] at a time t_2 given by

$$t_2 = \frac{3h^{\frac{1}{2}}(r)r^{\frac{1}{2}}}{h'(r)} = \frac{2h^{\frac{1}{2}}(r)}{3r^{\frac{3}{2}}\mu_0(r)} \equiv \frac{\sqrt{2M(r)}}{r^{\frac{3}{2}}\mu_0(r)}. \tag{12.23}$$

We note that

$$\frac{t_1}{t_2} = \frac{h'(r)r}{3h(r)}. \tag{12.24}$$

The monotony of μ_0 and the mean value theorem led Gu Chao Hao to the conclusion that

$$t_1 \leq t_2; \tag{12.25}$$

that is, the dust evolution does not induce a singularity in the metric before a dust shell arrives at the centre, in agreement with the previous conclusion of the absence of shell crossing. The first occurrence of the singularity is at a time t_0, with

$$t_0 = lim_{r=0}\left[\frac{r^3}{h(r)}\right]^{\frac{1}{2}} = \frac{1}{\sqrt{\frac{3}{2}\mu_0(0)}}; \tag{12.26}$$

the greater $\mu_0(0)$ is, the sooner the singularity appears.

12.2.2 *Matching with an exterior metric*

The interior metric is given by (12.19) for $r \leq a$, i.e.

$$0 \leq R \leq \{a^{\frac{3}{2}} - h^{\frac{1}{2}}(a)t\}^{\frac{2}{3}}. \tag{12.27}$$

The computations that we have made are still valid outside the star, where $r > a$ and $\mu_0(r) = 0$, but

$$M(r) = M_a := \int_0^a \mu_0(\rho)\rho^2 d\rho, \quad r > a. \tag{12.28}$$

The exterior metric reads

$$-\left(1 - \frac{2M_a}{R}\right)dt^2 + 2\sqrt{\frac{2M_a}{R}}dRdt + dR^2 + R^2(d\theta^2 + sin^2\theta d\phi^2). \tag{12.29}$$

This is a boosted Schwarzschild metric with horizon

$$R = 2M_a. \tag{12.30}$$

This solution takes the usual Schwarzschild form

$$-\left(1 - \frac{2M_a}{R}\right)dt^2 + \left(1 - \frac{2M_a}{R}\right)^{-1}dR^2 + R^2(d\theta^2 + sin^2\theta d\phi^2) \tag{12.31}$$

[20] Remark that a vanishing of R' signals a shell crossing.

if we make the change of time coordinate

$$d\tau = dt - \left(1 - \frac{2M_a}{R}\right)^{-1} \sqrt{\frac{2M_a}{R}} dR. \tag{12.32}$$

The full space time metric g is defined on $R^3 \times R^+$, with R, θ, ϕ polar coordinates on R^3 and $t \in R^+$ (i.e. $t \geq 0$) by

$$g = g_{dust}, \quad 0 \leq R \leq \{a^{\frac{3}{2}} - h^{\frac{1}{2}}(a)t\}^{\frac{2}{3}} \tag{12.33}$$

$$g = g_{ext}, \quad R \geq \{a^{\frac{3}{2}} - h^{\frac{1}{2}}(a)t\}^{\frac{2}{3}}. \tag{12.34}$$

The interior dust solution is hidden behind the horizon when

$$\{a^{\frac{3}{2}} - h^{\frac{1}{2}}(a)t\}^{\frac{2}{3}} < 2M_a. \tag{12.35}$$

This begins at a positive time t_3 given by

$$t_3 = h^{-\frac{1}{2}}(a)[a^{\frac{3}{2}} - (2M_a)^{\frac{3}{2}}], \tag{12.36}$$

if $a > 2M_a$; that is if the star is initially visible. We have seen that the outer shell of the star collapses into a singularity when $t_1(a) = h^{-\frac{1}{2}}(a)a^{\frac{3}{2}}$; hence after the star has ceased to be visible. This fact was an inspiration for the formulation of the **cosmic censorship conjecture** by Penrose.

We will return to the occurrence of singularities and cosmic censorship in Chapter 13.

13 The Reissner–Nordström solution

The **Reissner–Nordström** solution is a spherically symmetric solution of the Einstein–Maxwell equations. It describes the gravitational field of a charged spherically symmetric static body. It is found to be (Reissner, 1916), in standard coordinates,

$$-\left(1 - \frac{2m}{r} + \frac{Q^2}{2r^2}\right) dt^2 + \left(1 - \frac{2m}{r} + \frac{Q^2}{2r^2}\right)^{-1} dr^2 + r^2(d\theta^2 + sin^2\theta d\varphi^2). \tag{13.1}$$

The electromagnetic potential is

$$A_i = 0, \quad A_0 = -\frac{Q}{r}, \tag{13.2}$$

identical with the classical electrostatic potential of a spherical body with charge Q.

The metric (13.1) is smooth and Lorentzian with t a time variable as long as

$$\frac{2m}{r} - \frac{Q^2}{2r^2} < 1. \tag{13.3}$$

For large r the term Q/r^2 is small in comparison with m/r; since the total charge of celestial bodies is observed to be negligible in comparison to their

mass, the Reissner–Nordtröm solution has little application in astrophysics. It was also abandoned as a possible model for the electron, because for this particle it is estimated that $m^{-1}Q^2 = 2.8 \times 10^{-13}$ cm, and the influence of the term $2r^{-2}Q^2$ would be important only at distances where quantum effects cannot be neglected.

Though of little use in physics, the Reissner–Nordtröm solution is mathematically interesting[21]. Indeed, if $m^2 > Q^2$ it possesses two event horizons given in standard coordinates by

$$r_\pm = m \pm \sqrt{m^2 - Q^2}. \tag{13.4}$$

14 Schwarzschild spacetime in dimension n + 1

14.1 *Standard coordinates*

Reasonings and computations analogous to the ones made for $n + 1 = 4$ give the $n + 1$-dimensional Schwarzschild metric in spherical standard coordinates, $r \in R^+$, $d\omega^2$ the metric of the S^{n-1} sphere.

$$g_{Schw} = -\left(1 - \frac{2m}{r^{n-2}}\right) dt^2 + \left(1 - \frac{2m}{r^{n-2}}\right)^{-1} dr^2 + r^2 d\omega^2. \tag{14.1}$$

The Schwarzschild spacetime is defined for $n \geq 3$ by this metric supported by the manifold $M \times R$, with M the exterior of the ball $r^{n-2} = 2m$.

14.2 *Wave coordinates*

The same reasoning[22] as in the case $n = 3$ shows that the requirement that $x^\mu = (t, x^i)$ be wave coordinates, $\Box_g x^\mu = 0$, with $x^i = \bar{r}(r)n^i$, $n^i \in S^n$, reduces to the equation

$$\frac{d}{dr}\left[\frac{d\bar{r}}{dr} r^{n-1}(1 - 2mr^{2-n})\right] - (n-1)\bar{r} = 0 \tag{14.2}$$

(recall that if Δ^* is the Laplacian on the sphere S^{n-1}, then the considered functions x^i satisfy the equations $\Delta^* x^i = -(n-1)x^i$).

This equation is not equivalent to a Legendre equation if $n > 3$, as its solutions are not polynomials. Setting $s = 1/r$, one obtains an equation with a Fuchsian singularity at $s = 0$:

$$\frac{d}{ds}\left[s^{3-n}(1 - 2ms^{n-2})\frac{d\bar{r}}{ds}\right] = (n-1)s^{1-n}\bar{r}.$$

[21] It would invalidate a censorship conjecture if the word "generic" was not included in the hypotheses; see Chapter 13.

[22] Choquet-Bruat Y., Chruściel P., and Loiselet, J. (2006) *Class. Quant. Grav.*, **23**(24), 7383–94.

The characteristic exponents are -1 and $n-1$, so that, after matching a few leading coefficients, the standard theory of such equations provides solutions with the behaviour

$$\bar{r} = r + \frac{m}{(n-2)r^{n-3}} + \begin{cases} \frac{m^2}{4} r^{-3} \ln r + O(r^{-5} \ln r), & n = 4 \\ O(r^{5-2n}), & n \geq 5 \end{cases}$$

Somewhat surprisingly, there are logarithms of r in an asymptotic expansion of $\bar{r}$ in dimension $n = 4$. However, for $n \geq 5$ there is a complete expansion of $r - \bar{r}$ in terms of inverse powers of r, without any logarithmic terms. If we write g_{Schw} in the coordinates x^i, then

$$(g_{Schw,m})_{\mu\nu} = \eta_{\mu\nu} + (f_m)_{\mu\nu} \tag{14.3}$$

with the functions f of the form

$$f_{m,\mu\nu} = \frac{1}{r^{n-2}} h_{\mu\nu}\left(m, \frac{1}{r}, \frac{\bar{x}}{r}\right), \quad \bar{x} := (x^i), \tag{14.4}$$

with $f_{m,\mu\nu}(s, \bar{w})$ analytic functions of their arguments near $s = 0$, analytic in m and equal to zero for $m = 0$.

V

COSMOLOGY

1 Introduction

Due to the remarkable progress in astronomical observations, cosmology has recently become a full fledged part of physics.

A cosmological model is a spacetime which is supposed to represent the whole past, present, and future of our universe. It is argued that since the gravitational forces are long range (in contrast to nuclear forces) and non-compensated (there are no negative masses, though there are negative electric charges) these forces model the geometry of the cosmos. It is therefore assumed that the cosmological spacetime is a Lorentzian 4-manifold (V, g) which satisfies the Einstein equations with the non-gravitational energies as sources. In spite of the remarkable progress made in recent years with telescopes and satellites on observations of our cosmos, still a very small portion of it is known, and even there we have an imperfect knowledge of its energy content. Cosmological models therefore rely greatly on conjectures made *a priori* on philosophical prejudices, aiming for simplicity or based on the desire for unification. It is admitted by most cosmologists[1] that the cosmos is an Einsteinian spacetime at all scales, but with widely different representations for the sources, and even possibly with higher dimensions. Cosmologists implicitly admit, as experienced by everybody in daily life, that there is a flow of time, and though we do not observe the present state of the cosmos, since information comes to us at most with the speed of light, they speak of the Universe as being in this present state. Most cosmologists admit that the Universe is presently expanding; it started with the Big Bang, underwent a primeval phase where the modelling is at best speculative, and then an early phase – still speculative – where quantum phenomena were dominant, but where the physics was more like that which is accessible to experiment under extreme circumstances in our laboratories. To reconcile theory with different observations, some made at galactic scale, others at cosmological scale, modern cosmologists introduce new types of sources: dark matter and dark energy, also called quintessence.

In the classical models, still basic in cosmological studies, the cosmos is modelled by a four-dimensional Einsteinian spacetime with the source being a perfect fluid whose particles would be galaxies or clusters of galaxies. There is no evidence for the adequacy of such a representation, but other hypotheses, such as

[1] Though not quite by all of them; see the comments in Section 5.12.

a hierarchized or fractal Universe, do not seem, at present, to lead to tractable models.

The structure of the conjectured initial singularity (Big Bang or oscillatory) and the possible long-term future existence of the Universe are subjects of active mathematical research.

2 Cosmological principle

The **standard cosmological models** are based on the so-called **cosmological principle**, which is composed of two assumptions.

The first assumption is the existence of a family of **fundamental observers** in free fall. The timelines of the fundamental observers are geodesics which span the manifold V. Their proper time is called **cosmic time**.

The second important conjecture is that the Universe should look the same in all directions for any fundamental observer (isotropy) and the same to all observers (homogeneity). According to the fundamental definitions recalled in Chapter 2, the instantaneous space of an observer located at a point x of V is the hyperplane T_x orthogonal to the tangent vector u to its world line. It is meaningful to speak of homogeneity relative to the fundamental observers only if they have common space manifolds; that is, if their world lines are orthogonal to space sections, whose tangent plane is T_x. The manifold V is then a product $M \times R$. The lines $\{x\} \times R$ are timelike geodesics and the leaves $M_t = M \times \{t\}$ are spacelike; the spacetime metric can be written

$$^{(4)}g = -dt^2 + {}^{(3)}g \tag{2.1}$$

where t is the cosmic time and $^{(3)}g$ is a Riemannian metric on M, depending on t. A cosmos satisfying the cosmological principle is a Lorentzian manifold, $(M \times R, {}^{(4)}g)$, with a metric of the type (2.1), such that, for each t, the Riemannian manifold $(M, {}^{(3)}g)$, that is *the universe, is isotropic and homogeneous.*

The cosmological principle is an extension of the Copernican revolution which deprived our Earth of its central position. We also know that the Sun does not occupy a remarkable place in our galaxy. However, we see the stars very unevenly distributed in the night sky. Galaxies, and even clusters of galaxies, are also observed in our telescopes as very anisotropically and inhomogeneously distributed. The arguments in favour of adopting the cosmological principle are that, at a still larger scale, isotropy and homogeneity are attained. The strongest evidence for the validity of the cosmological principle is the existence of a cosmic background radiation which appears to us as very nearly isotropic. The discrepancy with isotropy is interpreted as due to our own motion[2] relative to the fundamental observers defined by the background radiation. An argument for homogeneity is that physical phenomena, in particular the fundamental constants, seem to be the same everywhere in the Universe we observe.

[2] This velocity is estimated to be about 200 km per second, i.e. very small with respect to the velocity of light.

3 Isotropic and homogeneous Riemannian manifolds

The more physically than mathematically minded reader can go directly to Section 4.

3.1 *Isotropy*

The isotropy of the space section M_t at a point x means that there is no privileged direction in the tangent space T_x to M_t at x. We give a mathematical formulation of isotropy in a Riemannian manifold, independent of its dimension, through the **sectional curvature**. The datum of the sectional curvature is equivalent to the datum of the Riemann tensor.

Definition 3.1

1. *The sectional curvature at a point x of a Riemannian manifold (M, g), relative to a 2-subplane P of the tangent space T_x, is the number, independent*[3] *of the choice of X and Y, linearly independent vectors in P,*

$$K(P) := \frac{Riemann(X, Y; X, Y)}{g(X, X)g(Y, Y) - g(X, Y)^2}; \tag{3.1}$$

that is, in local coordinates

$$K(P) \equiv \frac{R_{ij,hk} X^i Y^j X^h Y^k}{X^i X_i Y^j Y_j - (X^i Y_i)^2}. \tag{3.2}$$

2. *A Riemannian manifold (M, g) is said to be* ***isotropic*** *at x if its sectional curvature $K(P)$ at x is independent of the choice of the 2-plane P.*

Remark that the denominator in (3.1) is equal to 1 if X and Y are unit orthogonal vectors.

The following theorem characterizes isotropy in terms of the Riemann tensor.

Theorem 3.2 *(M, g) is isotropic at x if and only if its Riemann tensor at x takes the form, with $K(x)$ some number*

$$R_{ij,hk}(x) = K(x)(g_{ih}g_{jk} - g_{jh}g_{ik})(x). \tag{3.3}$$

Proof The definition 3.1 and the expression (3.3) imply that (M, g) has at x a sectional curvature independent of the 2-plane P, the number $K(x)$. The converse results from the algebraic properties of the Riemann tensor[4]. □

[3] The symmetry and antisymmetry properties of the Riemann tensor lead to the proof of the independence of $K(P)$ from the choice of (X, Y) in P (see for instance Kobayashi, S. and Nomizu, K. (1963) *Differential Geometry*, Vol. 1, Chapter 5, Prentice Hall.).

[4] See for instance Kobayashi, S. and Nomizu, K. (1969) *Riemannian Geometry*, Vol. 1, Chapter 5, Proposition 1.2.

3.2 *Homogeneity*

A Riemannian manifold (M, g) is called **homogeneous** if it admits a transitive[5] group of (global) isometries. It is called locally homogeneous if the group is only a pseudogroup of local isometries; that is, given two points x and y on M_t they admit isometric neighbourhoods.

Definition 3.3 *A Riemannian manifold (M, g) is said to be a* ***space of constant Riemannian curvature*** *if it is isotropic at each point $x \in M$ and the sectional curvature is a constant on M.*

The following theorem was proved by Schûr in 1886.

Theorem 3.4 *If a Riemannian manifold is isotropic at each point, it is a space of constant curvature.*

Proof If the Riemann tensor is given by (3.3), then the Ricci tensor and scalar curvature are

$$R_{ih} = (n-1)Kg_{ih}; \quad R = n(n-1)K. \tag{3.4}$$

Therefore the identity deduced from the Bianchi identity implies that

$$0 \equiv \nabla_h \left(R_i^h - \frac{1}{2}\delta_i^h R \right) = -\frac{1}{2}(n-1)(n-2)\partial_i K. \tag{3.5}$$

Therefore, if $n > 2$, K is a constant. □

Remark 3.5 *If g has constant curvature K, then an homothetic metric $\tilde{g} = kg$, k a constant, has constant curvature $\tilde{K} = k^{-1}K$.*

We now prove some important properties of Riemannian manifolds with constant Riemannian curvature. They rest on the following lemma.

Lemma 3.6 *Let (M, g) be a Riemannian space of constant curvature K. Then,*

1. *(M, g) is locally conformally flat.*
2. *In each domain of a chart there are local coordinates x^i in which the metric takes the form*

$$g = \frac{(dx^1)^2 + \cdots + (dx^n)^2}{1 + \frac{K}{4}[(x^1)^2 + \cdots + (x^n)^2]^2}. \tag{3.6}$$

Proof

1. Straightforward calculation shows that the Weyl tensor[6], if $n > 3$ (the Coton tensor if $n = 3$) of a metric with Riemann tensor of the form (3.3), vanishes identically.

[5] That is, for each pair of points x and y in M_t there exists an isometry of (M, g) which brings x onto y.

[6] See Appendix VI.

2. Since g is locally conformally flat, there are local coordinates in which the metric reads

$$g = \sum_{i=1,\dots,n} \frac{(dx^i)^2}{f^2(x^1,\dots,x^n)}. \tag{3.7}$$

Straightforward[7] calculation shows that the equations expressing that the Riemann tensor has constant curvature K reduce to

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = 0, \tag{3.8}$$

and

$$f\left[\frac{\partial^2 f}{(\partial x^i)^2} + \frac{\partial^2 f}{(\partial x^j)^2}\right] = K + \sum_h \left(\frac{\partial f}{\partial x^h}\right)^2, \quad i \neq j. \tag{3.9}$$

From the first set of equations it follows that

$$f = X_1 + \cdots + X_n \tag{3.10}$$

where X_i is a function of x^i alone. The second set of equations implies

$$\frac{\partial^2 f}{(\partial x^j)^2} = \frac{\partial^2 f}{(\partial x^k)^2} \quad \text{i.e.} \quad \frac{d^2 X_j}{(dx^j)^2} = \frac{d^2 X_k}{(dx^k)^2}. \tag{3.11}$$

Therefore, using the fact that a function of x^j can equal a function of x^k only if both are constant and equal, we find that

$$\frac{d^2 X_j}{(dx^j)^2} = c, \quad \text{hence} \quad \frac{dX_j}{dx^j} = cx^j + \lambda_j, \tag{3.12}$$

where λ_j is a constant which can be made zero by a change of coordinates. Then

$$X_j = \frac{c}{2}(x^j)^2 + \mu_j, \quad f = \frac{c}{2}\sum_i (x^i)^2 + \sum_i \mu_i \tag{3.13}$$

with μ_j a constant. We choose the local coordinates x^i such that $\sum_i \mu_i = 1$. Then

$$f = 1 + \frac{c}{2}\sum_i (x^i)^2. \tag{3.14}$$

Substituting this expression in (3.9) shows by a simple calculation that (3.9) is satisfied if and only if $K = 2c$, which gives the expression (3.6) of the metric g. □

We can reformulate point 2 of the lemma as the following theorem.

[7] See this type of calculation in Appendix VI.

Theorem 3.7 *A Riemannian manifold with constant curvature K is locally isometric to:*

If $K = 0$, the Euclidean space E^n.

If $K > 0$, a sphere of radius $K^{-\frac{1}{2}}$, submanifold

$$\sum_{A=1,\ldots,n+1} (x^A)^2 = K^{-1} \tag{3.15}$$

of the Euclidean space E^{n+1}.

If $K < 0$, a pseudo-sphere of pseudo-radius $|K|^{-\frac{1}{2}}$, also called hyperbolic space, the connected submanifold diffeomorphic to R^n,

$$\sum_{i=1,\ldots,n} (x^i)^2 - (x^{n+1})^2 = |K|^{-1}, \quad x^{n+1} > 0, \tag{3.16}$$

of the Minkowski space M^{n+1} of signature $(+\cdots+-)$.

Proof We leave the proof to the reader as an exercise in computing metrics induced on submanifolds. We will find a more familiar expression of Riemannian metrics with constant curvature in the next section. □

Corollary 3.8 *The spheres, Euclidean space, and pseudo-spheres are homogeneous spaces.*

Proof A Riemannian manifold with constant curvature K is locally homogeneous, by the definitions. The isometry groups of the respective flat spaces in which these manifolds are embedded induce transitive isometry groups on the embedded manifolds. □

Remark 3.9 *Isotropic at each point, but only locally homogeneous, spaces do not give spacetimes which satisfy in a strict sense the cosmological principle. This might seem a reason to reject them as cosmological models, but though isotropy of the cosmos (at a large scale) around us seems rather well confirmed, it is difficult to assert its global homogeneity. In fact some recent observations have led some cosmologists*[8] *to favour the multiple connectedness of our cosmos, which would imply, at best, the only local homogeneity of our Universe. We will return to the question of local homogeneity, regardless of isotropy, in a later section.*

4 Robertson–Walker spacetimes

We reserve the name **Robertson–Walker** for spacetimes which are the $3+1$-dimensional[9] models satisfying the cosmological principle; that is, to the metrics

[8] Luminet, J. P. *L'Univers Chiffoné*, Fayard.

[9] There is no difficulty in extending the definition to $n+1$ dimensions, but the topological and geometrical classification of three-dimensional manifolds do not extend to higher dimensions.

which read, on a product $M \times R$ with M a three-dimensional manifold,

$$^{(4)}g \equiv -dt^2 + {}^{(3)}\, g_t \tag{4.1}$$

where $^{(3)}g_t$ is a t-dependent metric of constant Riemann curvature, globally homogeneous. The manifold $(M, {}^{(3)}\, g_t)$ is referred to as "the **universe**".

4.1 *Space metrics*

If $^{(3)}g$ is of constant curvature, its Ricci tensor satisfies an equation

$$^{(3)}R_{ij} = 2K\ {}^{(3)}g_{ij}, \tag{4.2}$$

with K a constant. This equation is also sufficient for the constancy of the Riemann curvature of $^{(3)}g$, because in dimension 3 the Ricci tensor determines the Riemann tensor[10].

Due to their fundamental importance we give a different, simple, derivation of the local expressions of the space metrics $^{(3)}g$. We use the property of a space $(M, {}^{(3)}\, g)$ of constant curvature to be spherically symmetric around any of its points to write its metric in a local polar pseudo-coordinates neighbourhood (see Chapter 4) in the following form

$$^{(3)}g \equiv e^{\mu} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2), \quad \text{with} \quad \mu = \mu(r) \tag{4.3}$$

Equations (4.2) read then[11]

$$^{(3)}R_{ik} \equiv 0, i \neq k, \qquad {}^{(3)}R_{11} \equiv r^{-1}\mu' = 2Ke^{\mu}, \tag{4.4}$$

$$\sin^{-2}\theta\, {}^{(3)}R_{33} \equiv {}^{(3)}\ R_{22} \equiv -e^{-\mu} + 1 + \frac{r}{2}e^{-\mu}\mu' = 2Kr^2 \tag{4.5}$$

The general solution of (4.4) is trivially found to be $e^{-\mu} = -Kr^2$+constant; using (4.5) gives then:

$$e^{-\mu} = 1 - Kr^2. \tag{4.6}$$

We find again that a metric of constant curvature $K = 0$ is locally flat; a metric with constant curvature $K > 0$ [respectively $K < 0$] is locally a 3-sphere of radius $K^{\frac{1}{2}}$ [respectively locally a hyperbolic 3-space of radius $|K|^{-\frac{1}{2}}$]. When $K \neq 0$, we set $r = |K|^{-\frac{1}{2}}\bar{r}$. We can put the metric in one or other of the standard forms according to the sign of K (relabelling $\bar{r}$ as r)

$$^{(3)}g \equiv |K|^{-1}\gamma_{\varepsilon}, \quad \gamma_{\varepsilon} \equiv \frac{dr^2}{1 - \varepsilon r^2} + r^2(\sin^2\theta d\phi^2 + d\theta^2), \qquad \varepsilon = \text{sign}K. \tag{4.7}$$

[10] See the formula in Appendix VI.

[11] A prime is a derivative with respect to r. A dot denotes a derivative with respect to t.

The metric γ_ε is transformed into the familiar form of the unit sphere or pseudo-sphere metrics by setting

$$\text{When } \varepsilon = 1,\ r \equiv \sin\alpha,\ \text{then}$$

$$\gamma_+ = d\alpha^2 + \sin^2\alpha(\sin^2\theta d\phi^2 + d\theta^2) \tag{4.8}$$

$$\text{When } \varepsilon = -1,\ r = \sin h\chi,\ \text{then}$$

$$\gamma_- = d\chi^2 + sinh^2\chi(\sin^2\theta d\phi^2 + d\theta^2). \tag{4.9}$$

4.2 *Robertson–Walker spacetime metrics*

The previous section gives three types of Robertson–Walker spacetime metrics,

$$-dt^2 + R^2(t)\gamma_\varepsilon, \tag{4.10}$$

where R an arbitrary function of t. When $\varepsilon \neq 0$, $R(t)$ is equal to $|K|^{-\frac{1}{2}}$, K being the constant in space, time-dependent, Riemann curvature of ${}^{(3)}g$.

We see that $R(t)$ is a scaling factor of local spatial distances of fundamental observers. In an expanding universe, i.e. if $\dot{R} := dR/dt > 0$, this distance increases proportionally to R and to the original distance, like in an inflated balloon. Indeed on the trajectory of a fundamental observer only t varies; the space distance at time t of two observers moving respectively on lines $x = a$ and $x = b$ is the product by $R(t)$ of their distance in the metric γ_ε between these two points of M.

The metric corresponding to $\varepsilon = 1$ with $0 \leq \alpha < \pi$, (θ, ϕ) angular coordinates on S^2, is the canonical metric of the sphere S^3, a compact homogeneous Riemannian manifold. Universes with compact boundaryless spatial sections are sometimes called closed.

The metrics with $\varepsilon = 0$ or $\varepsilon = -1$ can both be supported as complete homogeneous metrics by the non-compact manifold R^3, with coordinates $0 \leq \chi < \infty$. The corresponding universes, spatially either the Euclidean space E^3 or the hyperbolic 3-space $\mathcal{H}^3$, are called open.

The manifolds S^3, R^3, and $\mathcal{H}^3$, are all simply connected and homogeneous.

4.3 *Robertson–Walker dynamics*

A Robertson–Walker (R–W) cosmology is a spacetime with a Robertson–Walker metric which satisfies the Einstein equations with sources

$$R_{\alpha\beta} = \Lambda g_{\alpha\beta} + \rho_{\alpha\beta} \equiv T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T. \tag{4.11}$$

The non-zero Christoffel symbols of the spacetime metric (4.10) are computed to be

$$\Gamma^0_{ij} \equiv R\dot{R}\gamma_{ij}, \quad \Gamma^j_{0i} \equiv R^{-1}\dot{R}\delta^j_i, \quad \Gamma^i_{jh} = \gamma^i_{jh}, \quad \dot{R} = \frac{dR}{dt}. \tag{4.12}$$

with γ^i_{jh} the Christoffel symbols for one or the other of the metrics γ_ε.

The non-zero components of the Ricci tensor of the spacetime metric are, using the value $r_{ij} = 2\varepsilon\gamma_{ij}$ for the Ricci tensor of the spatial metric γ_ε,

$$R_{00} \equiv -3R^{-1}\ddot{R}, \quad R_{ij} \equiv \{2\varepsilon + (R\ddot{R} + 2\dot{R}^2)\}\gamma_{ij} \tag{4.13}$$

We see that the symmetry required of a Robertson–Walker metric requires that the stress energy tensor of its sources has the algebraic form of a perfect fluid model; that is, such that

$$\rho_{\alpha\beta} \equiv T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T = (\mu + p)u_\alpha u_\beta + \frac{1}{2}g_{\alpha\beta}p, \tag{4.14}$$

where u is the fundamental observer's 4-velocity (non-tilted fluid in Wainwright–Ellis terminology), i.e. $u^0 = -u_0 = 1$, $u^i = u_i = 0$. The scalar functions μ and p are respectively the specific energy and pressure, depending only on t. The non-zero components of the tensor ρ are then

$$\rho_{00} \equiv \frac{1}{2}(\mu + 3p), \quad \rho_{ij} \equiv \frac{1}{2}R^2(\mu - p)\gamma_{ij}, \quad \text{hence } \gamma^{ij}\rho_{ij} \equiv \frac{3}{2}R^2(\mu - p). \tag{4.15}$$

The Einstein equations with cosmological constant Λ are

$$R_{00} = -\Lambda + \rho_{00}, \quad R_{ij} = \Lambda g_{ij} + \rho_{ij}$$

Modulo the identities (4.15), these equations read

$$-3R^{-1}\ddot{R} = \frac{1}{2}(\mu + 3p) - \Lambda, \tag{4.16}$$

$$2R^{-2}\varepsilon + (R^{-1}\ddot{R} + 2R^{-2}\dot{R}^2) = \frac{1}{2}(\mu - p) + \Lambda. \tag{4.17}$$

From these equations one deduces an expression for μ called the **Friedman equation**

$$\mu \equiv \frac{1}{2}\{R^{-2}\gamma^{ij}\rho_{ij} + \rho_{00}\} = 3R^{-2}(\dot{R}^2 + \varepsilon) - \Lambda, \tag{4.18}$$

and an expression for p

$$p = -2R^{-1}\ddot{R} - 2R^{-2}\varepsilon - R^{-2}\dot{R}^2 + \Lambda. \tag{4.19}$$

These equations imply that

$$\dot{\mu} = -3(\mu + p)R^{-1}\dot{R} \tag{4.20}$$

which can be rewritten

$$R\frac{d\mu}{dR} + 3\mu = -3p, \quad \text{i.e.} \quad \frac{d}{dR}(R^3\mu) = -3p. \tag{4.21}$$

Exercise. *Show that the relation reads as the conservation law* $\nabla_\alpha T^{\alpha 0} = 0$, *check that* $\nabla_\alpha T^{\alpha i} = 0$.

4.4 *Einstein static universe*

Historically, the first cosmological model was the **Einstein static universe**. It is a Robertson–Walker spacetime with $\varepsilon = +1$ and $R = R_0$, a constant; namely it is the manifold $S^3 \times R$ endowed with the static metric

$$-dt^2 + R_0^2 \gamma_+. \tag{4.22}$$

The components of its Ricci tensor are

$$R_{0\alpha} = 0, \quad R_{ij} = 2\gamma_{ij}$$

The Einstein static universe is a solution of Equations (4.18) and (4.19) with constant density and pressure and cosmological constant Λ:

$$\mu_0 = 3R_0^{-2} - \Lambda, \quad p_0 = -2R_0^{-2} + \Lambda \tag{4.23}$$

A negative pressure is unacceptable on classical physical grounds. This difficulty was remedied by Einstein by admitting a positive cosmological constant, an idea which he did not like, because it introduces a new parameter, *a priori* not geometrically defined.

A few years after introducing the cosmological constant (in order to obtain a time-independent cosmos) Einstein abandoned it (calling it the greatest blunder of his life) and accepted cosmological models which are time-dependent. In versions of these models where the universe is expanding, the observed red shift of light coming from distant galaxies is explained by the expansion of the Universe. Recent observations indicate that the cosmological expansion is accelerating. This has led some cosmologists to reintroduce a cosmological term in a time-varying form. The models for the early Universe led also to the introduction of a cosmological term. It is believed to be very small at present, and its variation with time is a puzzling subject.

The Einstein spacetime is, but with a different interpretation, the arena of the **Segal cosmos**: the manifold $S^3 \times R$ is the universal cover of the compactified Minkowski spacetime, and is conformally flat. The conformal group acts on this geometric, non-metric, structure. The representations of the conformal group govern the world of elementary particles. The cosmological red shift is explained by the existence of a local time different from the cosmic time.

4.5 *Cosmological red shift and the Hubble constant*

The essential cosmological data come from the observation of the red shift of stars and galaxies. The **red shift parameter** is defined to be

$$z \equiv \frac{\nu_S - \nu_0}{\nu_0}, \tag{4.24}$$

where ν_0 is the observed frequency and ν_S is the emitted frequency. It is a shift towards the red [respectively towards the blue] if $z > 0$ [respectively $z < 0$]. We now prove the following proposition.

Proposition 4.1 *In an expanding Robertson–Walker spacetime there is a cosmological red shift with parameter given by*

$$1 + z \equiv \frac{\nu_S}{\nu_0} = \frac{R(t_0)}{R(t_S)}, \tag{4.25}$$

if we assume that the variation of R is negligible during a period of the emission and reception of a light signal.

Proof Recall that an RW metric is of the form

$$-dt^2 + d\alpha^2 + f^2(\alpha)(d\theta^2 + \sin^2\theta d\phi^2) \tag{4.26}$$

Let O_0 and O_S be two fundamental observers. Take coordinates such that they have the same θ and ϕ and have α coordinates respectively 0 and α_S. Photons emitted by O_S at times t_S and $t_S + T_S$ follow null rays and arrive at O_0 at times t_0 and $t_0 + T_0$ which satisfy

$$\alpha_S = \int_{t_S}^{t_O} \frac{dt}{R(t)} = \int_{t_S+T_S}^{t_O+T_O} \frac{dt}{R(t)} \tag{4.27}$$

Elementary calculus shows that the times T_S and T_0 are therefore linked by the relation

$$\int_{t_S}^{t_S+T_S} \frac{dt}{R(t)} = \int_{t_O}^{t_O+T_O} \frac{dt}{R(t)}. \tag{4.28}$$

Hence if R is considered as constant during the small amount of time T_S (period of the emission) and T_O (period of the reception), we obtain

$$\frac{T_S}{T_O} = \frac{R(t_S)}{R(t_O)}. \tag{4.29}$$

Since the frequency is the inverse of the period this result leads to the desired relation (4.25). □

A positive cosmological red shift, i.e. $1 + z > 1$, signals an expansion of the Universe, $R(t_0) > R(t_S)$.

The statistical observations of distant galaxies confirm, in the accepted model, that our Universe is at present expanding. Most recent results seem to indicate that the expansion is accelerating.

In an R–W spacetime the distance of two given fundamental observers at some cosmic time t is proportional to $R(t)$. The **Hubble constant** is the constant in space, a t-dependent scalar

$$H \equiv R^{-1}\dot{R}. \tag{4.30}$$

It measures the rate of expansion (or possibly contraction) of the Universe. It has dimension $(\text{time})^{-1}$. An expanding Universe has (if Robertson–Walker) a positive Hubble constant.

The **Hubble law** says that the observed red shift is proportional to the distance of the source. It is a consequence of the theory only in a first approximation for not too distant sources, as we now show.

An observer O with spacetime coordinates $\alpha = 0$ (this is no restriction) and $t = t_0$ receives light from a spacetime point S with coordinates $\alpha = \alpha_S$, $t = t_S$ if it is on a light ray arriving at O. On this light ray it holds that $\theta =$ constant, $\phi =$ constant, and

$$\frac{dt}{d\alpha} = \pm R(t). \tag{4.31}$$

The signal left its source at time t_S if

$$\alpha_S = \int_{t_S}^{t_0} \frac{dt}{R(t)} \tag{4.32}$$

We see that since $R > 0$ and $t_S > t_O$, it follows from (4.32) that t_S (and consequently $R(t_S)$ as well) is a decreasing function of α_S. The red shift therefore increases with the distance of the source. More precise statements depend on the function $R(t)$.

In a first approximation, if $t_0 - t_S$ is small and if R is slowly varying, we have

$$\alpha_S \cong (t_0 - t_S)R^{-1}(t_S). \tag{4.33}$$

On the other hand, under the same hypothesis,

$$R(t_0) \cong R(t_S) + (t_O - t_S)\dot{R}(t_0). \tag{4.34}$$

Hence

$$z \cong dH(t_0) \quad \text{with} \quad d := \alpha_S R(t_0). \tag{4.35}$$

This *approximate* formula says that the red shift z is proportional to the present distance to the source.

A second approximation is obtained by introducing the **deceleration parameter**: it is a dimensionless scalar which measures the rate of variation with time of the expansion (or contraction), and is defined by

$$q = -\frac{R\ddot{R}}{\dot{R}^2}. \tag{4.36}$$

Applying the Taylor formula gives

$$R(t_S) = R(t_0)[1 + (t_S - t_0)H(t_0) - \frac{1}{2}qH^2(t_S - t_0)^2 + \cdots] \tag{4.37}$$

and

$$\alpha_S = R^{-1}(t_0)[t_0 - t_S + \frac{H}{2}(t_O - t_S)^2 + \cdots], \tag{4.38}$$

which gives the next approximation

$$z = Hd + \frac{1}{2}(q+1)H^2d^2 + \cdots . \tag{4.39}$$

4.6 *De Sitter spacetime*

De Sitter spacetime is one of the earliest cosmological found models which is not static. It is a vacuum RW Universe with a cosmological constant $\Lambda > 0$. It satisfies the equations

$$R_{00} = -\Lambda, \quad R_{ij} = \Lambda g_{ij}.$$

These equations reduce to

$$\ddot{R} - k^2 R = 0, \quad \text{with} \quad k^2 := \frac{\Lambda}{3}. \tag{4.40}$$

and

$$\dot{R}^2 - k^2 R^2 = -\varepsilon. \tag{4.41}$$

The general solution of (4.40) is, with A and B a pair of constants

$$R = Ae^{kt} + Be^{-kt}. \tag{4.42}$$

Such a function R verifies (4.41) in one or the other of the two cases:

1. $\varepsilon = 0$, either A or B equals zero. These solutions, defined on $R^3 \times R$, take the form

$$-dt^2 + e^{kt}\{dx^2 + dy^2 + dz^2\}. \tag{4.43}$$

   The spatial metric is Euclidean, up to a scale t-dependent factor; choosing $k > 0$, it crushes to zero as t tends to $-\infty$ while it expands to infinity as t tends to $+\infty$.

2. $\varepsilon = 1$, $4AB = k^{-2}$. In this case, the solutions are defined on $S^3 \times R$ with the metric of the form

$$g_{d\,\mathrm{Si}} := -dt^2 + k^{-2} cosh^2(kt)\{d\alpha^2 + \sin^2\alpha(d\theta^2 + \sin^2\theta d\phi^2)\}. \tag{4.44}$$

   The spacetime metric is time-symmetric. The spatial metric is the canonical metric of the sphere S^3 up to a scale t-dependent factor; it expands to infinity in both time directions.

This spacetime is called the **de Sitter** universe.

Theorem 4.2 *The de Sitter spacetime can be embedded in the flat five-dimensional space with Lorentzian metric*

$$-dT^2 + \sum_{A=1,\ldots,4} (dX^A)^2 \tag{4.45}$$

as the hyperboloid

$$T^2 - \sum_{A=1,\ldots,4} (X^A)^2 = k^{-2}, \qquad (\mathcal{H}). \tag{4.46}$$

Proof Define the embedding by

$$T = k^{-1}sh(kt), \quad X^A = k^{-1}ch(kt)\mu^A \tag{4.47}$$

where μ^A are parameters on S^3 as submanifold of E^4, for example

$$\mu^1 = \cos\alpha \quad \mu^2 = \sin\alpha\cos\theta, \quad \mu^3 = \sin\alpha\sin\theta\cos\phi, \quad \mu^4 = \sin\alpha\cos\theta\sin\phi. \tag{4.48}$$

The parameters μ^A are such that

$$\sum_{A=1,\ldots,4}(\mu^A)^2 = 1, \quad \text{hence} \quad \sum_{A=1,\ldots,4}\mu^A d\mu^A = 0. \tag{4.49}$$

On the hyperboloid it holds that

$$dT = T^{-1}\sum_{A=1,\ldots,4}X^A dX^A = cosh(kt)dt; \tag{4.50}$$

The metric induced from (4.45) on the hyperboloid is then found to be

$$\left[-dT^2 + \sum_{A=1,\ldots,4}(dX^A)^2\right]_{\mathcal{H}} = -dt^2 + k^{-2}cosh^2(kt)(dS^3)^2; \tag{4.51}$$

that is, the de Sitter metric (4.44). □

Remark 4.3 *Take on the half hyperboloid defined for $T + X^4 > 0$ coordinates $\tau, \xi^i \in R \times R^3$ by*

$$\tau := k^{-1}log[k(T + X^4)], \quad \xi^i = \frac{k^{-1}X^i}{T + X^4}. \tag{4.52}$$

The de Sitter metric reads in these coordinates

$$-d\tau^2 + e^{2k\tau}\sum_{i=1,2,3}(d\xi^i)^2. \tag{4.53}$$

The considered half of de Sitter spacetime is isometric to a spacetime with conformally Euclidean spacelike sections found in taking $\varepsilon = 0$.

Exercise. *Show that the de Sitter universe has constant spacetime curvature.*

Theorem 4.4 *The de Sitter spacetime is conformal to the slice $-\frac{\pi}{2} < t < \frac{\pi}{2}$ of the Einstein static universe.*

Proof Consider the diffeomorphism (equivalently change of coordinates) from de Sitter space time into the given slice of the Einstein universe, which is defined by

$$t' = 2\arctan(\exp kt). \tag{4.54}$$

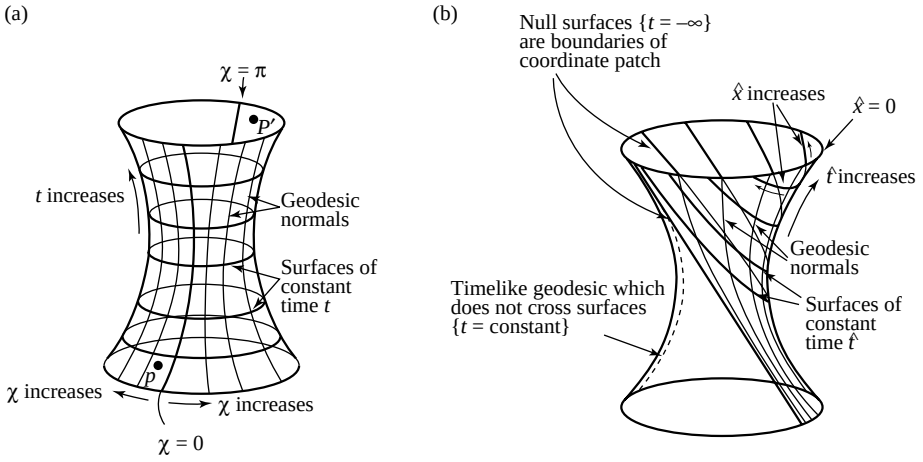


Figure 5.1 Representations of de Sitter spacetime.

An easy computation shows that

$$g_{d\,\mathrm{Si}} = k^{-2}(\cosh\chi)^2 g_{\mathrm{Ei}}. \tag{4.55}$$

□

De Sitter spacetime has been proved[12] to be the asymptotic state of Bianchi models (see Section 6) with positive cosmological constant, except the Bianchi type IX, for which Λ must be large enough for the statement to hold.

4.7 *Anti de Sitter (AdS) spacetime*

The anti de Sitter metric is a vacuum Robertson–Walker spacetime metric, but with negative cosmological constant. It satisfies the same equations (4.40) and (4.41) as the de Sitter metric time but the solution of (4.40) is now, setting $\Lambda = -3\omega^2$

$$R = A\cos\omega t + B\sin\omega t. \tag{4.56}$$

Metrics satisfying the Einstein equations (4.41) are obtained for $\varepsilon = -1$, in particular the spacetime metric

$$-dt^2 + \cos^2 t\gamma_-, \qquad \gamma_- := d\chi^2 + \sinh^2\chi(\sin^2\theta d\phi^2 + d\theta^2), \tag{4.57}$$

defined on $R^3 \times (-\frac{\pi}{2}, \frac{\pi}{2})$, is an Einsteinian vacuum spacetime with cosmological constant $\Lambda = -3$.

The spacetime classically called **anti de Sitter spacetime** is a static extension of such a spacetime. Specifically it is the manifold $R^3 \times R$ with metric

$$-dt^2(\cosh\chi)^2 + d\chi^2 + \sinh^2\chi(\sin^2\theta d\phi^2 + d\theta^2), \quad 0 \le \chi < \infty. \tag{4.58}$$

[12] Wald, R. (1984) *General Relativity*, University of Chicago Press.

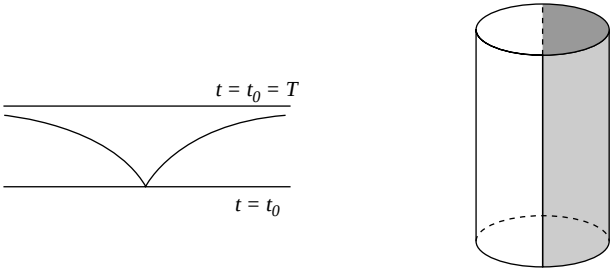


Figure 5.2 Light cone in AdS, conformal representation.

Remark that a radial light ray issued from a point with coordinates t_0, x_0 is such that

$$dt = \frac{d\chi}{\cosh\chi}, \quad \text{hence} \quad t = t^0 + \int_{\chi_0}^{\chi} \frac{d\chi}{\cosh\chi}. \tag{4.59}$$

We see that $t - t_0$ is bounded, whatever χ_0 is,

$$t - t_0 \leq \int_0^{\chi} \frac{d\chi}{\cosh\chi} < 2\int_0^{\infty} \frac{d\chi}{e^{\chi}} := T. \tag{4.60}$$

Therefore no light ray emitted at time t_0 will go beyond the space slice $t = t_0 + T$. The anti de Sitter space time has no global causality properties (it is not globally hyperbolic)[13], but it does play an important role in supergravity and string theory.

Theorem 4.5 *The anti de Sitter spacetime is conformal to the region* $0 \leq \alpha < \frac{\pi}{2}$ *of the Einstein cylinder*[14].

Proof The mapping $\alpha = 2\arctan(\exp\chi) - \frac{\pi}{2}$ is a diffeomorphism from R^3 (i.e. $\chi \in (0, \infty)$) onto $S^3 \cap \{0 \leq \alpha < \frac{\pi}{2}\}$. This diffeomorphism (i.e. change of coordinates) transforms the metric γ_- into $(\cosh\chi)^2\gamma_+$. □

5 Friedmann–Lemaître models.

These are models with source a perfect fluid.

5.1 *Equation of state*

Equation (4.21) gives a differential equation for R if an equation of state $p = p(\mu)$ is known. In the galactic epoch it is admitted that the energy is dominated by

[13] See Chapter 12.

[14] This property has been used (Choquet-Bruhat, Y. (1989) *Class. Quant. Grav.*, **6**, 1781–9.) to obtain the global existence of solutions of Yang–Mills equations on AdS spacetime with asymptotic boundary conditions.

the matter content and this is described by dust, i.e. $p = 0$. In the radiation-dominated era, i.e. in the early Universe, the equation of state is $p = \frac{1}{3}\mu$. Both cases can be treated simultaneously by taking the as equation of state the linear relation

$$p = (\gamma - 1)\mu. \tag{5.1}$$

One supposes in general that $\gamma \geq 1$, to have non-negative pressure for positive density; $\gamma = 1$ is the case of dust. The pure radiation case is $\gamma = \frac{4}{3}$. The sound wave velocity is less than the speed of light if $\gamma \leq 2$ (see Chapter 3).

5.2 *General properties*

Robertson–Walker universes without cosmological constant and with an equation of state of the type (5.1) are usually called **Friedmann–Lemaître** models.

For all values of γ, (5.1) and the energy equation (4.20) imply that

$$\dot{\mu} = -3\gamma\mu R^{-1}\dot{R}; \tag{5.2}$$

hence by integration, with M a constant,

$$\mu R^{3\gamma} = M. \tag{5.3}$$

Therefore the density μ must decrease with increasing R.

Using this relation reduces the Friedman equation (4.18) to a first-order differential equation for R:

$$\dot{R}^2 = \frac{1}{3}MR^{2-3\gamma} - \varepsilon, \tag{5.4}$$

whose resolution fixes the dynamics. Some qualitative features can be obtained with a few physical assumptions.

Beginning and age of the universe. For classical energy sources it holds that $\mu + 3p > 0$. Denote by t_0 the present time. By definition $R(t_0) > 0$; also $\dot{R}(t_0) > 0$, because we observe red shifts, not blue shifts (statistically). On the other hand the Raychauduri type equation (see Section XIII.3.3) together with the condition $\mu + 3p > 0$ imply that in all cases $\ddot{R} < 0$. Therefore the curve $t \mapsto R(t)$ is concave downwards. Since R is presently increasing, the curve $t \mapsto R(t)$ reaches the t-axis at some time less than t_0, which we take as $t = 0$ and consider as the beginning of our Universe (with the Big Bang). The tangent to the curve $t \mapsto R(t)$ at the present time t_0 intersects the t-axis at a point $T < 0$, such that, by elementary geometry

$$\dot{R}(t_0) = \frac{R(t_0)}{t_0 - T}. \tag{5.5}$$

Thus, by the definition of the Hubble constant $H(t)$, we have the following bound for the present age t_0 of the universe in cosmic time:

$$t_0 = H(t_0)^{-1} + T < H(t_0)^{-1}. \tag{5.6}$$

Behaviour with increasing t. This depends crucially on the value of ε. Indeed, the Friedmann equation implies the following results.

- $\varepsilon = -1$, $\dot{R}(t) \geq 1$, $R(t)$ keeps increasing at least linearly in t when t tends toward infinity.
- $\varepsilon = 0$, $R(t)$ also keeps increasing if $M > 0$, but more slowly.
- $\varepsilon = 1$, then R starts increasing and $R^{2-3\gamma}$ decreases since we have supposed $1 \leq \gamma \leq 2$. Then $\dot{R}$ vanishes when R attains the value $\sqrt[3\gamma-2]{\frac{M}{3}}$. That will be the end of the expansion; since $\ddot{R} < 0$ the scale factor R will start to decrease until it attains its big crunch at a finite time t_1.

5.3 *Friedmann models*

One usually calls Friedmann models those RW universes with space sections E^3, S^3, or $\mathcal{H}^3$ and with equation of state $p = 0$. Equation (5.4) reads then, with M a constant

$$\dot{R}^2 = \frac{1}{3} M R^{-1} - \varepsilon. \tag{5.7}$$

This can be integrated in terms of familiar functions:

- $\varepsilon = 0$. Taking $R(0) = 0$ gives

$$R(t) = c t^{\frac{2}{3}}, \quad c \equiv \left(\frac{3M}{4}\right)^{\frac{1}{3}} \tag{5.8}$$

and by scaling of r the metric reads

$$-dt^2 + t^{\frac{4}{3}}\{dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\}. \tag{5.9}$$

This spacetime is called the **Einstein–de Sitter universe** (Einstein and de Sitter, 1932).

- $\varepsilon = 1$. The reparametrization of t by T with

$$dT = \frac{dt}{R(t)} \tag{5.10}$$

leads to the solution

$$R = \frac{M}{6}(1 - \cos T) \quad \text{with} \quad t = \frac{M}{6}(T - \sin T). \tag{5.11}$$

Here t is an increasing function of T, while R begins increasing at the Big Bang $T = t = 0$, attains its maximum $\frac{M}{3}$ for $T = \pi$, i.e. $t = \frac{M\pi}{6}$, and then decreases to zero (big crunch) for $T = 2\pi$, i.e. $t = \frac{M\pi}{3}$.

- $\varepsilon = -1$. The same reparametrization of t leads to

$$R = \frac{M}{6}(\cosh T - 1). \tag{5.12}$$

The Universe starts expanding from the Big Bang at $t = T = 0$, and continues to expand forever, exponentially.

Asymptotic behaviour

In all three cases it holds that $t \cong \frac{MT^3}{6\times 6}$ when t approaches zero; hence in the neighbourhood of the initial Big Bang singularity it holds that

$$R(t) \cong \left(\frac{3M}{4}\right)^{\frac{1}{3}} t^{\frac{2}{3}}. \tag{5.13}$$

5.4 *Some other models*

5.4.1 *Flat Friedmann–Lemaître Universe*

Leaving $\varepsilon = 0$ with Euclidean space sections and equation of state (5.1) with $\gamma \geq 1$ and $p \geq 0$, we obtain the flat Friedmann–Lemaître models. The differential equation (4.4) for R reduces to

$$\dot{R}R^{\frac{3}{2}\gamma - 1} = C, \quad \text{a constant.} \tag{5.14}$$

Hence, with C_0 another constant, one has

$$R(t)^{\frac{3}{2}\gamma} = Ct + C_0. \tag{5.15}$$

With a careful choice of the origin of time and the scaling of r the metric reads

$$-dt^2 + t^{\frac{4}{3\gamma}}\{dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\}, \tag{5.16}$$

and for $\gamma = 1$ it reduces to the Einstein–de Sitter metric.

The Friedmann–Lemaître universe starts with a big bang at $t = 0$ and inflates indefinitely.

5.4.2 *Lemaître universe*

The **Lemaître** universe has $\varepsilon = 1$, space sections S^3, and a cosmological constant $\Lambda > 0$. It is not static.

5.4.3 *Milne universe*

The Milne universe has $\varepsilon = -1$, space support R^3, and is vacuum: $\mu = p = 0$. Hence

$$\dot{R}^2 = 1. \tag{5.17}$$

Up to the choice of label of the time origin, the spacetime metric on the manifold $R^3 \times (0, \infty)$ takes the form

$$-dt^2 + t^2\{dr^2 + sinh^2(r)(d\theta^2 + \sin^2(\theta)d\phi^2)\}. \tag{5.18}$$

where r, θ, ϕ are polar (pseudo-) coordinates on R^3. The space metric collapses for $t = 0$ and expands indefinitely when t tends to infinity. The spacetime metric is locally flat[15]. It is isometric to a wedge of Minkowski spacetime.

The **Milne universe**, or analogous ones with compact (non-simply connected) space sections, plays an important role as asymptotic states in cosmology.

[15] The Riemann tensor can be easily checked to be identically zero, for instance by using the $3 + 1$ decomposition given in Chapter 6.

5.5 *Confrontation with observations*

Fundamental observers of these mathematical model cosmologies do not readily compare to realistic physical observers. Neither our Earth revolving around the Sun, nor the Sun moving in our galaxy, nor any individual galaxy, follows the flow lines of the supposed cosmological fluid. Only the average of a great number of them (perhaps of clusters of clusters of them!) could be taken as following these trajectories[16]. Therefore confronting cosmological models with observation poses problems related to statistics.

The essential cosmological data, apart from the cosmological black body radiation, come from the observation of the red shift of stars and galaxies. In discussing cosmological observations it is useful to replace the density function μ by a dimensionless quantity Ω (the "**density parameter**") which is defined by

$$\Omega \equiv \frac{\mu}{3H^2}, \tag{5.19}$$

where $H := R^{-1}R'$ (the Hubble "constant"). Note that Ω depends only on cosmic time. The Friedmann equation (5.4) now reads

$$\Omega \equiv 1 + \frac{\varepsilon}{H^2R^2}. \tag{5.20}$$

If our universe is modelled by a Robertson–Walker cosmology, then the value of Ω determines the type of Robertson–Walker: $\Omega > 1$ implies that $\varepsilon = 1$, with closed space sections; $\Omega < 1$ implies $\varepsilon = -1$, with the spatial sections being open if they are simply connected; and $\Omega = 1$ (the "critical case" implies $\varepsilon = 0$, with Euclidean spatial sections, if they are simply connected.

Physicists argue that, in the classical models we are studying, if $\Omega > 1$ by an appreciable amount the Universe will collapse in an extremely short time. If $\Omega < 1$ it will expand so fast that no stars could form. Thus one must suppose that Ω is very close to 1. This estimate is in contradiction with the observation of visible matter. It has led to the introduction of "dark matter" and "dark energy", whose existences are only inferred from their theoretically predicted consequences.

6 Homogeneous non-isotropic cosmologies

The spaces associated with fundamental observers of a **homogeneous cosmology** admit a transitive Lie group of isometries, but their isotropy group at some point (hence at any point by transitivity), is not necessarily the full group $SO(3)$.

The isometry group G_r of an n-dimensional Riemannian manifold M is at most of dimension $r \leq \frac{1}{2}n(n+1)$. If G is transitive the orbit of any point x, i.e. the set of point images of x under isometries, is equal to M; then for any x and

[16] It was pointed out to me by T. Damour that the velocity of the Milky Way, and of the Earth, with respect to the cosmological black body radiation is comparatively small, therefore it is, in a first approximation, legitimate to assimilate us with fundamental observers defined through this fundamental observational fact.

y in M there exists an element $T \in G$ such that $Tx = y$, and therefore $r \geq n$. The group G_r is said to be simply transitive if each point is attained only once; then $r = n$. If $r > n$, then at a point, and hence at each point by transitivity, the manifold (M, g) admits a non-trivial isotropy group of dimension $s = r - n$[17].

Since we consider the physical case $n = 3$, the isotropy group has dimension $s = 0, 1$ or 3 (there are no two-dimensional subgroups of the rotation group of E^3). The case $s = 3$ corresponds to homogeneous and isotropic models. The spatially homogeneous anisotropic models are called **Bianchi models**. The name "Bianchi model" is extended to spacetimes which are only locally homogeneous. Bianchi models are said to be **locally rotationally symmetric (LRS)** if they have, in addition, an isotropy group of dimension 1.

Bianchi classification

To a Lie group G is associated a Lie algebra $\mathcal{G}$, realized by the vector space of the tangent vector fields to G which are invariant under the left translations of G, endowed with an associative anticommutative multiplication called the Lie bracket, defined through the Lie derivative $\mathcal{L}$ by

$$[X, Y] = -[Y, X] := \mathcal{L}_X Y - \mathcal{L}_Y X. \tag{6.1}$$

The Lie bracket satisfies the Jacobi identity: if X, Y, Z are elements of $\mathcal{G}$ then

$$[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0. \tag{6.2}$$

If $\mathcal{G}$ is finite dimensional (which is the case under study) the Lie bracket is determined by **structure constants** c^i_{jh}. Choosing a basis $X_{(i)}$ on $\mathcal{G}$ they are such that

$$[X_{(j)}, X_{(h)}] = c^i_{jh} X_{(i)}. \tag{6.3}$$

Under a change of basis the structure constants behave tensorially on $\mathcal{G}$; they are antisymmetric in the lower indices and satisfy the Jacobi identity deduced from (6.2).

Conversely, for a given Lie algebra, there are associated several topologically inequivalent Lie groups[18], but there is one and only one which is connected and simply connected[19].

The Bianchi cosmologies are classified along with the Lie algebras of their three-dimensional transitive Lie groups of spatial isometries. The classification of three-dimensional Lie algebras is due to Bianchi. Bianchi cosmologies have been studied extensively by Taub[20] (1951).

[17] This is proven by showing that one can construct $r - d$ Killing vector fields which vanish at x.

[18] For example R^n and T^n have the same, trivial, Lie algebra.

[19] See for instance CB-DM1.

[20] Reprinted in Taub, A. H. (2004) *General Relativity and Gravitation*, **36**(12), 1699–719. See also the book by Ryan, M. and Shepley, L. (1975) *Homogeneous Relativistic Cosmologies*, Princeton University Press.

One defines on $\mathcal{G}$ the covariant vector A to be the trace

$$A_i := c^j_{ji}, \tag{6.4}$$

and the 2-tensor Q to be given by

$$Q^{ij} := \frac{1}{2}\varepsilon^{ihk}(c^j_{hk} - \delta^j_h A_k) \tag{6.5}$$

with ε^{ihk} the canonical[21] totally antisymmetric 3-tensor.

Lemma 6.1

1. *Q is a symmetric 2-tensor.*
2. *Q together with A satisfy*

$$Q^{ij}A_i = 0. \tag{6.6}$$

3. *The structure constants are given in terms of Q and A by*

$$c^j_{lm} = \varepsilon_{ilm}Q^{ij} + \frac{1}{2}(\delta^j_l A_m - \delta^j_m A_l). \tag{6.7}$$

Proof

1. Computing for instance Q^{12} using the definition (6.5) gives

$$Q^{12} = -\frac{1}{2}(c^1_{13} + c^2_{23}) = Q^{21}. \tag{6.8}$$

2. Use the Jacobi identity.
3. Follows from the definitions. □

The **Bianchi classification** is done using the rank and signature of the quadratic form Q. Keeping the traditional (though not quite logical), names it reads as follows:

- **Bianchi class A**. The vector A is zero. If $\mathcal{G}$ is the Lie algebra of an Abelian group or of a compact group, or the direct product of two such groups, then it is included in this class.

  $Q \equiv 0$, i.e. abelian Lie algebra: Bianchi class I.
  Q of rank 1: Bianchi class II
  Q of rank 2, signature $+\ -$ Bianchi class VI_0
  Q of rank 2, signature $+\ +$ Bianchi class VII_0
  Q of rank 3, signature $++\ -$ Bianchi class VIII
  Q of rank 3, signature $+\ +\ +$ Bianchi class IX.

[21] That is such that $\varepsilon^{123} = \varepsilon_{123} = 1$.

- **Bianchi class B**, the vector A is different from 0. The group is not compact. The quadratic form Q admits zero as an eigenvalue hence is of rank at most 2.

The Bianchi classes B are named as follows.
$Q \equiv 0$, Bianchi class V.
Q of rank 1: Bianchi class IV
Q of rank 2, signature $+\ -$ Bianchi class VI_h
Q of rank 2, signature $+\ +$ Bianchi class VII_h
The "missing" Bianchi type III is identical to Bianchi VI_{-1}

As noted above, a Lie algebra determines a local Lie group. Several non-isomorphic Lie groups can have the same Lie algebras.

7 Bianchi class I universes

The Robertson–Walker spacetimes with $\varepsilon = 0$ are a particular family of homogeneous (in this case also isotropic) Bianchi class I universes.

7.1 *Kasner solutions*

The **Kasner spacetimes** were discovered by Kasner (1925), then studied as cosmological models by Lemaître (1933) and by Heckmann and Shucking (1962)[22]. They have attracted considerable interest for the study of the behaviour of spacetimes near the initial singularity, since the work of Belinski *et al.* (1970)[23].

The Kasner models are built with the isometry group G being the Abelian group R^3; they are of Bianchi class I. All structure constants are zero.

Let x^i be arbitrary Cartesian coordinates on R^3; the differentials dx^i are a basis of invariant 1-forms on R^3. We can choose them at each time t so that they are orthogonal in the metric g of the corresponding orbit. This metric then takes the diagonal form

$$g \equiv \sum_{i=1,2,3} a_i(t)(dx^i)^2. \tag{7.1}$$

The vacuum Einstein equations are ordinary differential equations which read

$$^{(4)}R_{00} \equiv -\frac{1}{4}(\partial_0 \log a_i)^2 - \frac{1}{2}\partial_0\partial_0 \log a_i = 0 \tag{7.2}$$

$$^{(4)}R_{ij} \equiv 0 \quad \text{if} \quad i \neq j \tag{7.3}$$

$$^{(4)}R_{ii} \equiv -\left\{\frac{1}{2}(\partial_0 \log a_i)\partial_0 a_i - \frac{1}{4}\partial_0 a_i \sum_{p=1,2,3}(\partial_0 \log a_p) - \frac{1}{2}\partial_{00}^2 a_i\right\} = 0 \tag{7.4}$$

[22] Heckmann, O. and Shucking, E. (1962) in *Gravitation, an Introduction to Current Research* (ed. L. Witten), Wiley.
[23] Belinski, V. A., Kalatnikov, I. M., and Lifshitz, E. M. (1970) *Adv. Phys.*, **19** 525–73.

We set

$$\frac{1}{2}\partial_0 \log a_i = b_i. \tag{7.5}$$

Then (7.2) reads

$${}^{(4)}R_{00} \equiv \sum_{i=1,2,3} \{{b_i}^2 + \partial_0 b_i\} = 0. \tag{7.6}$$

Using the elementary calculus identity

$$\frac{1}{2}a_i^{-1}\partial_{00}^2 a_i \equiv \partial_0\left(\frac{1}{2}a_i^{-1}\partial_0 a_i\right) + \frac{1}{2}a_i^{-2}(\partial_0 a_i)^2, \tag{7.7}$$

we find that (7.4) implies

$$a_i^{-1}{}^{(4)}R_{ii} \equiv b_i\left(\sum_{j=1,2,3} b_j\right) + \partial_0 b_i = 0, \tag{7.8}$$

and hence it follows that

$$g^{ij}{}^{(4)}R_{ij} \equiv \left(\sum_{i=1,2,3} b_i\right)^2 + \sum(\partial_0 b_i) = 0 \tag{7.9}$$

The two equations (7.6) and (7.9) imply the constraint

$${}^{(4)}S_{00} \equiv \frac{1}{2}({}^{(4)}R_{00} + g^{ij}{}^{(4)}R_{ij}) \equiv -\left(\sum_{i=1,2,3} b_i\right)^2 + \sum_{i=1,2,3} b_i^2 = 0. \tag{7.10}$$

Letting K denote the extrinsic mean curvature of the space sections $t =$ constant, we calculate

$$K \equiv -\sum_{i=1,2,3} b_i \equiv -X^{-1}\partial_0 X \tag{7.11}$$

with

$$X := (\det g)^{\frac{1}{2}} \equiv (a_1 a_2 a_3)^{\frac{1}{2}}. \tag{7.12}$$

Equation (7.7) gives

$$\partial_0 K = K^2, \quad \text{equivalently} \quad \partial_{00}^2 X = 0. \tag{7.13}$$

The solution whose volume tends to zero when t tends to zero (correspondingly K becomes infinite) takes the form

$$X = t, \quad K = -\frac{1}{t}. \tag{7.14}$$

Equations (7.8) then become a diagonal system of first-order differential equations for the functions b_i, namely

$$ {}^{(4)}R_{ii}a_i^{-1} \equiv \frac{b_i}{t} + \partial_0 b_i = 0. \tag{7.15} $$

The general solution of this equation becomes infinite for $t = 0$. It takes the form, with p_i a constant,

$$ b_i = \frac{p_i}{t}. \tag{7.16} $$

Hence

$$ a_i = t^{2p_i}. \tag{7.17} $$

The **Kasner exponents** p_i must verify, due to (7.10) and (7.14)

$$ \left(\sum_{i=1,2,3} p_i \right)^2 = \sum_{i=1,2,3} p_i^2 = 1. \tag{7.18} $$

The vacuum Einstein equations are then all satisfied. The Kasner spacetime metric is, explicitly

$$ -dt^2 + t^{2p_1}(dx^1)^2 + t^{2p_2}(dx^2)^2 + t^{2p_2}(dx^3)^2 \tag{7.19} $$

where the Kasner exponents p_i lie in the Kasner circle (7.18), the intersection of a 2-sphere and a plane.

One of the Kasner solutions has two of the Kasner exponents vanishing. In this case, the spacetime metric is locally flat, as can be checked easily by computing the Ricci tensor of the 2-metric

$$ -dt^2 + t^2 dx^2. \tag{7.20} $$

The spacetime with such a Kasner metric supported by the manifold $R^3 \times R^+$ is in fact isometric to the wedge $t > |x|$ of the Minkowski spacetime.

Exercise. *Find this isometry.*

Hint[24]*. Use the change of coordinates*

$$ t = (t'^2 - x'^2)^{\frac{1}{2}} \qquad x' = \tanh^{-1}(\frac{x'}{t'}). \tag{7.21} $$

Asymptotic properties

For all Kasner solutions the volume of M_t expands from zero to infinity as t increases from zero to infinity, since $X := (\det g)^{\frac{1}{2}} = t$.

If two of the exponents are not zero the relation (7.18) shows that one at least must be negative. Suppose $p_1 < 0$, $p_2 > 0$, and $p_3 > 0$. Then as t tends to zero the spacetime shrinks in the direction of x^2 and x^3 while it expands indefinitely in the direction of x^1. The opposite happens as t tends to infinity; in both time directions the Universe is very anisotropic (cigar shaped), while it is much less so at intermediate times.

[24] See for instance Wald (1984), Section 6.4. Exchanging the roles of t and x^1 gives the **Rindler spacetime.**

7.2 *Models with matter*

We now consider the case of Bianchi type I with solutions containing comoving (non-"tilted") dust. The stress energy tensor is given by

$$\rho_{00} = \frac{1}{2}\mu, \quad \rho_{ij} = \frac{1}{2}\mu g_{ij}, \quad \rho_{0i} = 0, \tag{7.22}$$

with μ a function of time only.

The constraints then reduce to the single equation

$${}^{(4)}S_{00} \equiv -\left(\sum_{i=1,2,3} b_i\right)^2 + \sum_{i=1,2,3} b_i^2 = \mu. \tag{7.23}$$

The conservation equation governing the motion of matter, $\nabla_\alpha(\mu u^\alpha) = 0$, reads

$$\partial_0(\mu X) = 0 \quad \text{with} \quad X \equiv (a_1 a_2 a_3)^{\frac{1}{2}}; \tag{7.24}$$

it implies, with an integration constant which it is convenient to call $\frac{3}{2}m$

$$\mu = \frac{3}{2}\frac{m}{X}. \tag{7.25}$$

Equation (7.9) is now replaced by

$$g^{ij}{}^{(4)}R_{ij} \equiv (\partial_0 Y)^2 + \partial_{00}^2 Y = \frac{9}{2}me^{-Y}, \tag{7.26}$$

where we have set $X = e^Y$, so that $\sum b_i = \partial_0 Y$. this can be re-expressed as

$$\partial_{00}^2 e^Y - \frac{9}{2}m = 0 \tag{7.27}$$

which gives

$$e^Y \equiv X = 9\frac{m}{2}t(t+k) \tag{7.28}$$

with k a constant and X vanishing for $t = 0$. When X is known the equations ${}^{(4)}R_{ii} = \rho_{ii}$ become, as in the vacuum case, an ordinary differential diagonal system for the a_is. They read

$$a_i^{-1}{}^{(4)}R_{ii} \equiv b_i\frac{\partial_0 X}{X} + \partial_0 b_i = \frac{1}{2}\mu = \frac{3m}{4X}; \tag{7.29}$$

i.e.

$$\partial_0(b_i X) = 3\frac{m}{4}. \tag{7.30}$$

The b_i must satisfy the constraint (7.23), which reads

$${}^{(4)}S_{00} \equiv -\left(\sum_{i=1,2,3} b_i\right)^2 + \sum_{i=1,2,3} b_i^2 = \frac{3}{2}\frac{m}{X}. \tag{7.31}$$

The a_i must satisfy the condition

$$a_1 a_2 a_3 = X^2. \tag{7.32}$$

Combining these results one finds (see Hawking and Ellis (1973)[25] and references therein) a solution of the form

$$a_i = X^{\frac{2}{3}} (t^{\frac{2}{3}} X^{-\frac{1}{3}})^{2p_i} \tag{7.33}$$

with $X(t)$ is given by (7.28) and the numbers p_i are given by

$$p_1 = 2\sin\alpha, \quad p_2 = 2\sin\left(\alpha + \frac{2\pi}{3}\right), \quad p_3 = 2\sin\left(\alpha + \frac{4\pi}{3}\right) \tag{7.34}$$

for some constant α.

Exercise. *Show that the given solution satisfies all the Einstein equations.*

Like the vacuum Kasner spacetime, this solution exhibits anisotropy and different rates of expansion in the different directions, since the p_i are of different signs. However, it tends to be isotropic for large t. Indeed for $t \to \infty$

$$X \cong t^2 \tag{7.35}$$

$$a_i \cong t^{\frac{4}{3}}, \tag{7.36}$$

up to a multiplicative constant; hence the Universe tends to be isotropic.

On the other hand, near $t = 0$, up to a multiplicative constant, if $k \neq 0$ then

$$X \cong tk \tag{7.37}$$

$$a_i = (tk)^{\frac{2}{3}} (t^{\frac{1}{3}} k^{-\frac{1}{3}})^{2p_i} \tag{7.38}$$

which gives anisotropic behaviour near $t = 0$. In fact the allowed values of the p_i require that one (but not all) of them is negative; some direction is therefore expanding while the others are collapsing.

This spacetime is thus a model in which early anisotropy is smoothed into isotropy by evolution due to the presence of matter.

8 Bianchi type IX

The structure constants of a type IX Lie algebra are antisymmetric in all their indices, with

$$c^3_{12} = 1. \tag{8.1}$$

[25] Hawking, S. and Ellis, G. (1973) *The Large Scale Structure of Spacetime*, Cambridge University Press.

The corresponding simply connected Lie group is $SU(2)$, identified with the sphere S^3. A basis of right invariant 1-forms on S^3 takes the following form in Euler coordinates (see for instance CB-DM1, Problem III 5)

$$\theta^1 \equiv \sin\psi d\theta - \cos\psi \sin\theta d\phi, \quad \theta^2 \equiv \cos\psi d\theta + \sin\psi \sin\theta d\phi \tag{8.2}$$

$$\theta^3 \equiv \cos\theta d\phi + d\psi, \quad 0 \leq \psi \leq 4\pi, \quad 0 \leq \phi \leq 2\pi, \quad 0 \leq \theta \leq \pi \tag{8.3}$$

Exercise. *Check that* $d\theta^1 = -\theta^2 \wedge \theta^3$ *and analogous relations for* $d\theta^i$, $i = 2, 3$.

We consider a Bianchi type IX universe with a diagonal metric of the form

$$-dt^2 + \sum_i a_i(t)(\theta^i)^2 \tag{8.4}$$

The vacuum Einstein equations in vacuum ${}^{(4)}R_{00} = 0$ and ${}^{(4)}R_{ij} = 0$ are the same as (7.2) and (7.3), except for the addition to ${}^{(4)}R_{ij}$ of R_{ij}, the Ricci tensor of the 3-metric. Specifically, with $b_i \equiv \frac{1}{2}\partial_0 \log a_i$, one has

$${}^{(4)}R_{00} \equiv \sum_i \{{b_i}^2 + \partial_0 b_i\} \equiv \sum_i \frac{\partial_{00}^2 a_i^{\frac{1}{2}}}{a_i^{\frac{1}{2}}} = 0 \tag{8.5}$$

and with $X^2 \equiv a_1 a_2 a_3$,

$$a_i^{-1}{}^{(4)}R_{ii} \equiv b_i X^{-1}\partial_0 X + \partial_0 b_i + a_i^{-1} R_{ii} = 0. \tag{8.6}$$

A straightforward computation shows that this differential equation can be written, with a circular permutation for the other indices 2 and 3, in the form

$$X^{-1}\partial_0(a_2^{\frac{1}{2}} a_3^{\frac{1}{2}} \partial_0 a_1^{\frac{1}{2}}) = -a_{11}^{-1} R_{11} \tag{8.7}$$

with

$$a_{11}^{-1} R_{11} \equiv \frac{1}{2X^2}\{(a_1)^2 - (a_2 - a_3)^2\}, \tag{8.8}$$

One does not obtain (as in the Kasner case) a differential equation for X whose solution permits the writing of a decoupled differential system for the a_i's. However, following Belinski, Kalatnikov, and Lifshitz (1970) one can speculate as follows on the asymptotic behaviour of the spacetimes:

Asymptotic properties

Equations (8.5) and (8.6) are identical to the corresponding Kasner equations if their right-hand sides are equated to zero. In that case we have the Kasner solution

$$a_i = t^{2p_i}, \quad \text{with} \quad \sum_i p_i = \sum_i p_i^2 = 1. \tag{8.9}$$

Suppose there is a range of values of t for which the considered equations have a solution given approximately by

$$a_i \cong t^{2p_i} \tag{8.10}$$

Equation (8.5) then gives

$$\sum_i p_i = \sum_i p_i^2. \tag{8.11}$$

This equation implies (the trivial case of all p_is zero being excluded) that $\sum_i p_i > 0$, hence a volume expanding with t increasing from 0 to infinity.

Elementary algebra shows that (8.9) implies that not all of the p_is are positive. Suppose $p_1 < 0$, $0 < p_2 \leq p_3$. Then, for small t, a_2 and a_3 can be neglected in comparison with a_1. One introduces a new time variable τ defined by

$$\frac{dt}{d\tau} = X, \tag{8.12}$$

and sets

$$a_i^{\frac{1}{2}} \equiv e^{\alpha_i} \tag{8.13}$$

Assuming the hypotheses we have made, the functions α_i satisfy approximately the following differential equations (a prime denotes a derivative with respect to τ)

$$\alpha_1'' + \frac{1}{2} e^{4\alpha_1} \cong 0 \tag{8.14}$$

$$\alpha_2'' = \alpha_3'' \cong \frac{1}{2} e^{4\alpha_1} \tag{8.15}$$

Equation (8.14) shows that α_1 will be an oscillating function of τ. It will start increasing after some amount of τ time and the approximation will cease to be valid. Belinski, Lifshitz and Kalatnikov [BKL]. Introduce the so called Kasner eras, which are a succession of approximations by Kasner solutions during which the three axes alternatively shrink and expand, with Kasner exponents on the Kasner circle jumping in a random way. These heuristic considerations have given birth to many interesting studies on the possible chaotic behaviour of the Universe in the vicinity of the Big Bang. The most recent and complete study, for general spacetimes in arbitrary dimension, done by Damour and collaborators, using the Hamiltonian and the walls it creates[26], is reviewed in a section contributed by T. Damour in Chapter 13.

9 The Kantowski–Sachs models

It is proven that any group G_4 admits a subgoup G_3. If the transitive isometry group G_4 admits a subgroup G_3 which acts multitransitively on M, its orbits have necessarily[27] dimension $s = 2$. These 2-spaces have constant curvature K,

[26] and an Iwasawa decomposition of the metric, $g_{ij} = \sum_{a=1}^{d} e^{-2\beta^a} N_i^a N_j^a$ with N an upper triangular matrix with 1s on the diagonal.

[27] To have a G_3 of isometries.

positive, zero, or negative. The homogeneous spacetimes with such a transitive group G_4 are called **Kantowski–Sachs** models, their metrics taking the following form:

$$-dt^2 + A^2(t)dx^2 + B^2(t)g^2(x)(d\theta^2 + f^2(\theta)d\phi^2) \tag{9.1}$$

with $f(\theta) = \sin\theta$ $(K > 0)$, $f(\theta) = \theta$ $(K = 0)$, $f(\theta) = \sinh\theta$ $(K < 0)$.

It was proved[28] that the case $K > 0$ is the only one in which the transitive group G_4 does not admit a simply transitive G_3. The Bianchi types of the groups G_3 are IX $(K > 0)$, VII_0 $(K = 0)$, or VIII $(K < 0)$.

10 Taub and Taub NUT spacetimes

10.1 *Taub spacetime*

In 1951 Taub discovered a spatially homogeneous vacuum spacetime with topology $S^3 \times R$. We describe this space now.

A manifold M diffeomorphic to the sphere S^3 has a principal fibre bundle structure with group and fibre S^1 and base diffeomorphic to S^2, a particular case is the Hopf fibreing (see for instance CB-DM2, p. 307), the metric of S^3 is invariant under the action of the group. In the Euler coordinates of S^3 a left invariant 1-form on the fibre is $d\psi$, and the projection on the basis is given by $(\psi, \theta, \phi) \mapsto (\theta, \phi)$, with (θ, ϕ) coordinates on $\Sigma \cong S^2$. The manifold $V = R \times M$ has also a principal fibre bundle structure with group and fibre S^1 and base $R \times \Sigma$, $\Sigma \cong S^2$. A metric on V invariant under the S^1 action reads (see Appendix VII)

$$\Phi^2(d\psi + \alpha)^2 + {}^{(3)}g \tag{10.1}$$

with Φ a function, α a locally defined 1-form and ${}^{(3)}g$ a Lorentzian metric on $\Sigma \times R$, The manifold of the **Taub spacetime** spacetime is the above fibre bundle V. The metrics ${}^{(3)}g$ induce on each $R \times \{x\}$, $x \in \Sigma$, a metric independent of x and on $\Sigma \times \{t\}$, up to a factor depending only on t, the metric of S^2, invariant by a transitive three-parameter group. The space manifolds $(M_t \cong S^3, g_t)$ of the Taub spacetime are therefore homogeneous, in fact invariant by a four-parameter group, which belongs to type IX of the Bianchi classification and is moreover Locally Rotationally Symmetric (abbreviated to **LRS** in the usual classifications).

Taub proves that the following choice of Φ, α and ${}^{(3)}g$ gives a solution of the Einstein equations in Euler coordinates on S^3, which admits the indicated symmetries.

$$-U^{-1}dt^2 + (2\ell)^2 U(d\psi + \cos\theta d\phi)^2 + (t^2 + \ell^2)(d\theta^2 + \sin^2\theta d\phi^2) \tag{10.2}$$

where, with ℓ and m positive constants,

$$U \equiv -1 + \frac{2(mt + \ell^2)}{t^2 + \ell^2} \tag{10.3}$$

[28] See Ryan, M. and Shepley, L. (1975) *Homogeneous Relativistic Cosmologies*, Princeton University Press.

The Taub spacetime is regular and globally hyperbolic on $S^3 \times (t_-, t_+)$ where

$$t_\pm = m \pm (m^2 + \ell^2)^{\frac{1}{2}} \tag{10.4}$$

values of t for which U vanishes. It is timelike and null geodesically incomplete.

10.2 *Taub NUT spacetime*

It has been found by Newman, Unti, and Tamburino that the Taub spacetime can be extended, though as a non-globally hyperbolic spacetime.

Indeed the change of coordinates

$$\psi' = \psi + \frac{1}{2\ell} \int_0^t \frac{d\tau}{U(\tau)} \tag{10.5}$$

makes the singularity disappear due to the power -1 of U in the Taub metric, which now reads as a C^∞ metric on $S^3 \times R$

$$\begin{aligned} &4\ell^2 d\psi'^2 - 4\ell d\psi' dt + 4\ell^2 U \cos\theta d\phi d\psi' - 2\ell \cos\theta dt d\phi \\ &\quad + (t^2 + \ell^2)(d\theta^2 + \sin^2\theta d\phi^2) \end{aligned} \tag{10.6}$$

The two surfaces, diffeomorphic to S^3, $t = m \pm (m^2 + \ell^2)^{\frac{1}{2}}$, which bound the Taub spacetime in the **Taub-NUT**, are null surfaces; they are the Cauchy horizons of the maximal development of any Cauchy surface in the Taub spacetime. They are generated by closed null geodesics where ψ only varies (fibres of the Hopf fibreing).

One family of null geodesics issued from a point of the Taub region crosses both horizons; another family spirals near these surfaces and is incomplete. In the NUT region the fibres are closed timelike curves (the surfaces $t = constant^t$ are there timelike).

Another extension of the Taub spacetime with analogous properties, though not isometric to it, is obtained by the change of coordinates

$$\psi'' = \psi - \frac{1}{2\ell} \int_0^t \frac{d\tau}{U(\tau)} \tag{10.7}$$

Both these Taub NUT spaces are incomplete.

The Taub NUT spaces give counterexamples to several conjectures. The consolation is that they are not generic, due to their symmetries.

11 Locally homogeneous models

We have already shown that for a universe $(V, {}^{(4)}g)$ with V diffeomorphic to $M \times R$ and metric of the form (2.1), the isotropy at each point of ${}^{(3)}g$ implies that ${}^{(3)}g$ has constant curvature, hence is locally homogeneous; but it does not imply that $(M, {}^{(3}g)$ is globally homogeneous. We have also remarked that, although isotropy of the cosmos (at a large scale) around us seems rather well confirmed, it is difficult to assert its global homogeneity, and even its topology. The relations

between the topology and the possible metric structures of compact Riemannian manifolds have remarkable properties, particularly in dimension 3, which we briefly survey[29]. The manifolds and metrics we consider are supposed to be smooth.

Remark 11.1 *Every topological manifold of dimension less than or equal to 3 has a differentiable structure, unique up to diffeomorphisms close to the identity. There exist topological manifolds of dimension greater than 3 which have none, or many, differentiable structures. The sphere S^7 has 28, S^{11} has 992. The spaces R^n have one, except R^4 which, as a topological manifold, has an infinity of differentiable structures.*

11.1 *n-dimensional compact manifolds*

We first give some properties which do not depend on the dimension n of the manifold.

We denote by $I_g(M)$ the isometry group of the Riemannian manifold (M, g) and by I_g^0 its connected component of the identity.

The following definition is independent of the choice of a particular Riemannian metric on M; it is a topological invariant.

The **degree of symmetry** of a manifold M is the non-negative integer

$$deg\ M := max \text{ dimension of } I_g M, \quad g \text{ any Riemannian metric on } M. \tag{11.1}$$

One knows that $deg\ M \leq \frac{1}{2}n(n+1)$.

Theorem 11.2 *Let M be a compact manifold. Then the following are equivalent*

1. *$deg\ M = 0$;*
2. *For each metric g on M, the isometry group $I_g(M)$ is finite;*
3. *No metric g on M admits a Killing vector field.*

Proof $2 \rightarrow 1$. If $I_g(M)$ is finite, it has dimension zero, hence if $I_g(M)$ is finite for all metrics g on M, then $deg\ M = 0$.

$1 \rightarrow 2$. If $degM = 0$ the isometry group $I_g(M)$ of any manifold (M, g) is discrete; since it is also compact if M is compact, it is finite.

$2 \rightarrow 3$ and $3 \rightarrow 2$. When M is compact the Lie algebra of the isometry group $I_g(M)$ of (M, g) is the Lie algebra of its Killing vector fields. $I_g(M)$ is finite if (M, g) admits no Killing vector fields[30]. □

The following easy to prove theorem implies, in particular, that compact manifolds of constant negative Riemann curvature can be only locally homogeneous.

[29] This section is largely inspired by Fisher, A. and Moncrief, V. (1994) in *Global Structures and Evolution in General Relativity* (eds. S. Cotsakis and G. Gibbons), pp. 111–73. Springer-Verlag, London.

[30] The zero vector field is not considered as a Killing vector field.

Theorem 11.3 *Compact Riemannian manifolds with negative Ricci curvature do not admit global Killing vector fields.*

Proof Killing vector fields satisfy the equation

$$\nabla_i v_j + \nabla_j v_i = 0. \tag{11.2}$$

Using the Ricci identity and the equation $\nabla_i v^i = 0$, consequence of the previous one, we obtain

$$\nabla^i(\nabla_i v_j + \nabla_j v_i) = \nabla^i \nabla_i v_j + R^i_j v_i = 0. \tag{11.3}$$

Contracting this equation with v, integrating on the compact manifold M, and then integrating by parts leads to

$$\int_M \{-|\nabla v|^2 + R_{ij} v^i v^j\} d\mu_g = 0 \tag{11.4}$$

This equation implies that $v \equiv 0$ if the quadratic form $Ricci(g)$ is negative definite. It implies that v has vanishing covariant derivative if $Ricci(g) \equiv 0$. □

The following properties are enjoyed by compact $n-$dimensional manifolds. Their proofs are often quite involved.

Definition 11.4

1. *A complete Riemannian manifold (M, g) of constant curvature K is called a **spherical space form** if $K > 0$, a **Euclidean space form** if $K = 0$, or a **hyperbolic space form** if $K < 0$.*

2. *A manifold M is called of **flat type** (respectively, of **spherical type, hyperbolic type**) if there exists on M a metric g such that (M, g) is a Euclidean space form, (respectively, a spherical space form, a hyperbolic space form)*

It is known[31] that manifolds of flat type are diffeomorphic to a quotient R^n/Γ where Γ is a subgroup of the Euclidean group acting freely and properly discontinuously on R^n; manifolds of hyperbolic type are diffeomorphic to quotients $\mathcal{H}^n/\Gamma$ where Γ is a subgroup of the isometry group of $\mathcal{H}^n$ acting freely and properly discontinuously on $\mathcal{H}^n$.

In both of the cases listed above the universal cover of the manifold is diffeomorphic to R^n. One defines:

Definition 11.5 *A $K(\pi, 1)$ manifold is a manifold whose universal cover is contractible*[32].

A $K(\pi, 1)$ manifold M is also called **aspherical** because any continuous map of a sphere S^m, $m > 1$, into M is homotopic to a constant map, i.e. the image can be continuously deformed to a point.

[31] See references in Fisher and Moncrief (1994).

[32] A manifold is contractible if it can be continuously deformed to a point. R^n is a contractible manifold; it is not the only one, but other examples are complicated.

11.2 *Compact 3-manifolds*

It is particularly interesting for cosmology to know the relations between the topology of a 3-manifold with the kind of Riemannian metrics it can support, especially for the locally homogeneous metrics.

It has been known since the 19th century, in particular from the work of Poincaré, that every orientable compact surface – i.e. 2-dimensional manifold – admits a metric with constant curvature $+1$, 0 or -1, and hence is locally isometric to S^2, R^2, or $\mathcal{H}^2$; in fact it is diffeomorphic to the quotient of one of these manifolds by a discrete isometry group. This property permits a classification of orientable compact surfaces as "cushions with holes", the number of holes is equal to the numbers of cuts one must inflict to the surface to make it simply connected, it is called the **genus** of the surface. Topology and metric structure are related through the Euler–Poincaré characteristic and the Gauss–Bonnet formula. Such a classification is not possible in dimension greater or equal to 4, because of the complexity of the fundamental groups. In dimension 3, there is no *a priori* obstruction to the classification.

Riemannian manifolds of dimension 3 with zero Ricci curvature, *a fortiori* Riemannian manifolds with constant curvature zero, are locally flat[33]. The 3-torus T^3 is a compact locally flat space, and is the only one (up to homothetie, i.e. change of scale) which is (globally) homogeneous. There are five other non-diffeomorphic, locally flat 3-manifolds, with isometric degree equal to 1 or zero.

The Poincare conjecture, now proved by Perelman, following previous works, in particular by Hamilton, states that all compact simply connected[34] – in other words with a trivial fundamental group[35] – 3-manifolds are diffeomorphic[36] to the sphere S^3.

Modulo the conjecture, generally believed to be true, that the only free action of finite groups on S^3 are $SO(4)$ actions, it is known that

Proposition 11.6 *(Elliptization conjecture) Every compact orientable 3-manifold with finite fundamental group is diffeomorphic to a spherical space form; that is, can support a metric of constant positive curvature.*

This proposition is part of the **Thurston classification conjecture** which posits a relationship between the topology and geometry of compact and oriented 3-dimensional manifolds (for Riemannian geometers a geometric structure on a differentiable manifold M is a complete locally homogeneous Riemannian metric g). We refer the interested reader to the specialized mathematical literature for further information.

[33] That is, locally isometric to Euclidean space.
[34] That is, such that any closed curve can be continuously deformed to a single point.
[35] The homotopy group of closed loops.
[36] Speaking here of diffeomorphisms is equivalent to speaking of homeomorphisms due to the previous remark.

In the final section on cosmology, we return briefly to physics, also encouraging the interested reader to go to specialized texts for further reading.

12 Recent observations and conjectures

A combination of supernova data and measurements of the microwave background anisotropies strongly suggests for Friedman cosmology the flat model, $\varepsilon = 0$, dominated by vacuum energy.

The existence of dark matter (non-luminous objects) is established by the observation that various luminous objects – stars, galaxies or even clusters of galaxies – move faster than one would expect if they only felt the attraction of visible objects. The possible particles corresponding to dark matter are called WIMPS (Weakly Interacting Massive Particles). The recently asserted positivity of the neutrino mass is not sufficient to account for the observed effects.

To explain the strong differences between the state of the Universe as it is conjectured at the Big Bang and its present stage, and to explain the formation of cosmological structures starting from small gravitational instabilities, most cosmologists rely now on inflation, a kind of phase transition linked with a scalar field, called initially the inflaton. This scalar field Φ would satisfy a wave equation with a potential $V(\Phi)$, whose variation could explain the different structures of the Universe at early times and now. This potential would have in the beginning a positive quasi-plateau, interpreted in equations as a positive cosmological constant, and would account for inflation. It would now be near its almost zero minimum, corresponding to an almost zero cosmological constant[37], as seems to be the result of different independent categories of observations[38]. A scalar field, called dark energy, or quintessence, could also solve some other gaps between theory and observations.

Dark energy and dark matter have not yet been observed in laboratories.

Remark, however, that in cosmological models issued from unified field theories, in particular from string theory, the Universe we observe is obtained by a Kaluza–Klein reduction. Such a reduction results in the appearance, in addition to the usual Maxwell and Yang Mills fields, of a multiplet of scalar fields.

Even accepting the postulate, which has up to now been able to fit observational data, that our cosmos is an Einsteinian spacetime, one cannot help to remark that many non Friedmann models, including inhomogeneities, can be contemplated to model our cosmos. In fact, astronomers seem always to speak of the universe as a three-dimensional object which evolves with time. This notion is in contrast with the general spirit of Relativity, where time and space cannot be globally separated. The notion of such a final modelling of the cosmos is even more in contrast with more wild speculations, sometimes inspired by facts which

[37] About 10^{-23} in geometrical units according to recent estimates.

[38] Physicists observe that, if the cosmological "constant" had not taken those values, we would not be here to study it: there would have been no atom formation in the first period and no galaxy formation in sufficient numbers in the second period for observers to exist.

we could not even conceive before, like quantum mechanics and its possible many universes interpretation.

In conclusion, cosmology is a fascinating subject, which leaves open many questions, but may help us not only to have a greater knowledge of the Universe we live in, but also a better understanding of fundamental physical laws.

VI

LOCAL CAUCHY PROBLEM

1 Introduction

Since we are treating non-quantum fields, it seems reasonable that we consider the problem of classical dynamics, i.e. the problem of the evolution of initial data. The Einstein equations are a geometrical system, invariant under diffeomorphisms of V, the associated isometries of g, and transformation of the sources. From the analyst's point of view they constitute, for the metric, a system of second-order quasilinear partial differential equations. However, the system is underdetermined because of the Bianchi identities, a consequence of the diffeomorphism invariance: the equations are not independent and their characteristic determinant is identically zero.

The system is overdetermined, and we will see that data on a submanifold are not independent.

In the Cauchy problem for the Einstein equations, space and time play different roles. We first give the general $n+1$ splitting of these equations formulated in terms of time-dependent space tensors, through the use of an arbitrary moving frame. The natural frame of vectors tangent to coordinate curves is a particular case of a moving frame. We will find the canonical decomposition of a sliced Lorentzian metric carried out in Chapter 1 particularly useful.

2 Moving frame formulae

2.1 *Frame and coframe*

A moving frame in a subset U of a differentiable $\mathrm{n}+1$-dimensional manifolds V is a set of $n+1$ vector fields on U which are linearly independent in the tangent space T_xV at each point $x \in U$. A coframe on U is a set of $\mathrm{n}+1$ 1-forms θ^α which are linearly independent at each $x \in U$ in the dual space T_x^*V. In the domain U of a coordinate chart a coframe is specified by $n+1$ linearly independent differential 1-forms

$$\theta^\alpha \equiv a^\alpha_\beta dx^\beta, \tag{2.1}$$

with a^α_β functions on U.

Remark 2.1 *The product structure $V = M \times R$, with M an orientable 3-manifold, implies the existence of a global coframe (but not of global coordinates!). The property does not extend to higher dimensions.*

The dual frame to a coframe θ^α is the set of vector fields such that

$$\theta^\alpha(e_\beta) = \delta^\alpha_\beta. \tag{2.2}$$

The dual of the natural frame $\frac{\partial}{\partial x^\alpha}$ is the coframe defined by the 1-forms $\theta^\alpha \equiv dx^\alpha$. In this case, $d\theta^\alpha \equiv 0$. More generally, the differentials of the 1-forms θ^α do not vanish; they are given by 2-forms

$$d\theta^\alpha \equiv -\frac{1}{2} C^\alpha_{\beta\gamma} \theta^\beta \wedge \theta^\gamma \tag{2.3}$$

The functions $C^\alpha_{\beta\gamma}$ on U are called the **structure coefficients** of the frame.

We note that the basis e_α dual to θ^α then satisfies the commutation conditions

$$[e_\alpha, e_\beta] = \frac{1}{2} C^\gamma_{\alpha\beta} e^\gamma,$$

where $[.,]$ denotes the Lie bracket of vector fields[1].

The Pfaffian derivative ∂_α in the coframe θ^α of a function f on U is defined by

$$df \equiv \frac{\partial f}{\partial x^\alpha} dx^\alpha \equiv \partial_\alpha f \ \theta^\alpha. \tag{2.4}$$

If we denote by A, with elements A^β_α, the matrix inverse of a with elements a^β_α (see Equation (2.1), we have

$$\partial_\alpha f \equiv A^\beta_\alpha \frac{\partial f}{\partial x^\beta}. \tag{2.5}$$

Pfaffian derivatives do not commute. One deduces from (2.3) and from the identity $d^2 f \equiv 0$ that

$$d^2 f \equiv \frac{1}{2} \{\partial_\beta \partial_\gamma f - \partial_\gamma \partial_\beta f - C^\alpha_{\beta\gamma} \partial_\alpha f\} \theta^\beta \wedge \theta^\gamma \equiv 0; \tag{2.6}$$

hence

$$(\partial_\alpha \partial_\beta - \partial_\beta \partial_\alpha) f \equiv C^\gamma_{\alpha\beta} \partial_\gamma f. \tag{2.7}$$

2.2 *Metric*

A metric on U is a non degenerate quadratic form of the θ^α,

$$g \equiv g_{\alpha\beta} \theta^\alpha \theta^\beta. \tag{2.8}$$

The metric is **Lorentzian** if the quadratic form is of Lorentzian signature, $(-+++\cdots+)$.

A frame is called orthonormal for the metric g if $g_{\alpha\beta} = \pm 1$. In the case of a Lorentzian metric we denote by θ^0 the timelike (co)axis and θ^i the space (co)axis. Then, with respect to an orthonormal frame, $g_{00} = -1$ and $g_{ij} = \delta_{ij}$, the Kronecker symbol.

[1] See for instance CB-DM1.

2.3 *Connection*

A linear connection on V permits the definition of an intrinsic derivation of vectors and tensors[2]. It is specified on the domain U by a matrix-valued 1-form ω; i.e. by a set of matrix-valued 1-forms ω^β_γ linked to functions $\omega^\beta_{\alpha\gamma}$ by the identities

$$\omega^\beta_\gamma := \omega^\beta_{\alpha\gamma}\theta^\alpha. \tag{2.9}$$

The covariant derivative of a vector v with components v^α is

$$\nabla_\alpha v^\beta := \partial_\alpha v^\beta + \omega^\beta_{\alpha\gamma} v^\gamma \tag{2.10}$$

An analogous formula holds for a covariant vector, but with a minus sign in front of $\omega^\beta_{\alpha\gamma}$.

The connection ω is called the **Riemannian connection** of g if

- It has vanishing torsion; i.e.

$$d\theta^\gamma + \omega^\gamma_{\alpha\beta}\theta^\alpha \wedge \theta^\beta = 0. \tag{2.11}$$

  That is

$$\omega^\alpha_{\beta\gamma} - \omega^\alpha_{\gamma\beta} = C^\alpha_{\beta\gamma}$$

  The interpretation of this condition is that the second covariant derivatives of scalar functions commute, namely

$$\nabla_\alpha \partial_\beta f - \nabla_\beta \partial_\alpha f \equiv 0.$$

  In particular, in the natural frame, $\omega^\alpha_{\beta\gamma}$ is symmetric in β and γ.
- The covariant derivative of the metric is zero; i.e.

$$\partial_\alpha g_{\beta\gamma} - \omega^\lambda_{\alpha\gamma} g_{\beta\lambda} - \omega^\lambda_{\alpha\beta} g_{\lambda\gamma} = 0. \tag{2.12}$$

The conditions (2.11) and (2.12) imply by straightforward computation that

$$\omega^\beta_{\alpha\gamma} \equiv \Gamma^\beta_{\alpha\gamma} + g^{\beta\mu}\tilde{\omega}_{\alpha\gamma,\mu} \tag{2.13}$$

with

$$\tilde{\omega}_{\alpha\gamma,\mu} \equiv \frac{1}{2}(g_{\mu\lambda}C^\lambda_{\alpha\gamma} - g_{\lambda\gamma}C^\lambda_{\alpha\mu} - g_{\alpha\lambda}C^\lambda_{\gamma\mu}) \tag{2.14}$$

$$\Gamma^\beta_{\alpha\gamma} \equiv \frac{1}{2} g^{\beta\mu}(\partial_\alpha g_{\gamma\mu} + \partial_\gamma g_{\alpha\mu} - \partial_\mu g_{\alpha\gamma}). \tag{2.15}$$

The quantities Γ are called the Christoffel symbols of the metric g. The connection coefficients reduce to them in the natural frame. They are zero for an orthonormal frame, and, more generally, when the $g_{\alpha\beta}$ are constant.

[2] See CB-DM1, p. 300.

2.4 *Curvature*

2.4.1 *Definition*

The non-commutativity of covariant derivatives is a geometrical property of the metric, via its Riemannian connection. The curvature measures this non-commutativity.

The Riemann curvature tensor is an exterior 2-form taking values in the set of linear maps from the tangent plane to itself. It is defined by[3],

$$(\nabla_\lambda \nabla_\mu - \nabla_\mu \nabla_\lambda) v^\alpha \equiv R_{\lambda\mu,}{}^{\alpha}{}_{\beta} v^\beta, \tag{2.16}$$

from which we calculate

$$R_{\lambda\mu,}{}^{\alpha}{}_{\beta} \equiv \partial_\lambda \omega^\alpha_{\mu\beta} - \partial_\mu \omega^\alpha_{\lambda\beta} + \omega^\alpha_{\lambda\rho}\omega^\rho_{\mu\beta} - \omega^\alpha_{\mu\rho}\omega^\rho_{\lambda\beta} - \omega^\alpha_{\rho\beta} C^\rho_{\lambda\mu}. \tag{2.17}$$

2.4.2 *Symmetries and antisymmetries*

The Riemann tensor is a symmetric double 2-form: it is antisymmetric in its first two indices, and in its last two indices written in covariant form. It is invariant under the interchange of these two pairs. This last property is a consequence of the identity

$$R_{\alpha\beta,\lambda\mu} + R_{\mu\alpha,\beta\lambda} + R_{\lambda\mu,\alpha\beta} \equiv 0. \tag{2.18}$$

2.4.3 *Bianchi identities*

The definition (2.16) implies the identity (vanishing of the covariant differential of the curvature 2-form):

$$\nabla_\alpha R_{\beta\gamma,\lambda\mu} + \nabla_\beta R_{\gamma\alpha,\lambda\mu} + \nabla_\gamma R_{\alpha\beta,\lambda\mu} \equiv 0. \tag{2.19}$$

2.4.4 *Ricci tensor, scalar curvature, Einstein tensor*

The **Ricci tensor** is defined by

$$R_{\alpha\beta} := R_{\lambda\alpha,}{}^{\lambda}{}_{\beta}. \tag{2.20}$$

The **scalar curvature** is

$$R := g^{\alpha\beta} R_{\alpha\beta}. \tag{2.21}$$

The **Einstein tensor** is

$$S_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R. \tag{2.22}$$

2.4.5 *Conservation identity*

Contracting the Bianchi identities implies the identity

$$\nabla_\alpha R_{\beta\gamma,}{}^{\alpha}{}_{\mu} - \nabla_\beta R_{\gamma\mu} + \nabla_\gamma R_{\beta\mu} \equiv 0; \tag{2.23}$$

[3] CB-DM1 p. 306.

and a further contraction gives the following identity satisfied by the Einstein tensor:

$$\nabla_\alpha S^{\alpha\beta} \equiv 0. \tag{2.24}$$

This identity implies that the source must satisfy the so called **conservation laws**

$$\nabla_\alpha T^{\alpha\beta} = 0 \tag{2.25}$$

if the metric satisfies the Einstein equations with source $T_{\alpha\beta}$:

$$S_{\alpha\beta} = T_{\alpha\beta}. \tag{2.26}$$

3 $n+1$ splitting adapted to space slices

3.1 *Adapted frame and coframe*

We consider a spacetime with manifold $V = M \times R$ and hyperbolic metric g such that the submanifolds $M_t \equiv M \times \{t\}$ are spacelike. We take a frame with space axes e_i tangent to the space slice M_t and time axis e_0 orthogonal to it. Such a frame, particularly adapted to the solution of the Cauchy problem, is called a **Cauchy adapted frame**. We take local coordinates adapted to the product structure, $(x^\alpha) = (x^i, x^0 = t)$, and we choose for e_i the vectors $\partial/\partial x^i$ of a natural frame on M_t, i.e.

$$\partial_i = \frac{\partial}{\partial x^i} \tag{3.1}$$

The dual coframe is found to be such that, with β a time-dependent vector tangent to M_t called the **shift**,

$$\theta^i = dx^i + \beta^i dt, \tag{3.2}$$

while the 1-form θ^0 does not contain dx^i. We choose

$$\theta^0 = dt. \tag{3.3}$$

The vector e_0, i.e. the Pfaffian derivatives ∂_0, is then

$$\partial_0 \equiv \partial_t - \beta^j \partial_j \quad \text{with} \quad \partial_t := \partial/\partial t. \tag{3.4}$$

The vector e_0 is timelike since it is orthogonal to spacelike surfaces. We suppose that it defines the positive time orientation. In the coframe θ^α one has $g_{i0} = 0$, and the metric reads

$$ds^2 = -N^2(\theta^0)^2 + g_{ij}\theta^i\theta^j, \quad \theta^0 = dt, \quad \theta^i = dx^i + \beta^i dt. \tag{3.5}$$

The function N is called the **lapse**. We shall assume throughout that $N > 0$. The time-dependent, properly Riemannian, space metric induced by g on M_t is denoted either g_t or $\bar{g}$. An overbar denotes a spatial tensor or operator, i.e. a t-dependent tensor or operator on M. Note that, in our frame, $\bar{g}_{ij} = g_{ij}$ and $\bar{g}^{ij} = g^{ij}$.

A spacetime (V, g) with $V = M \times R$ and metric (3.5) is called a **sliced spacetime**.

3.2 *Structure coefficients*

The structure coefficients of a Cauchy adapted frame are found to be

$$C^i_{0j} = -C^i_{j0} = \partial_j \beta^i; \tag{3.6}$$

all other structure coefficients are zero.

3.3 *Splitting of the connection.*

We denote by $\bar{\nabla}$ the covariant derivative corresponding to the space metric $\bar{g}$. Using the general formulae (2.12–14) we find that

$$\omega^i_{jk} \equiv \Gamma^i_{jk} \equiv \bar{\Gamma}^i_{jk} \tag{3.7}$$

$$\omega^i_{00} \equiv N g^{ij} \partial_j N, \quad \omega^0_{0i} \equiv \omega^0_{i0} \equiv N^{-1} \partial_i N, \quad \omega^0_{00} \equiv N^{-1} \partial_0 N,$$

and

$$\omega^0_{ij} \equiv \frac{1}{2} N^{-2} \{\partial_0 g_{ij} + g_{hj} C^h_{i0} + g_{ih} C^h_{j0}), \tag{3.8}$$

from which we obtain

$$\omega^0_{ij} \equiv \frac{1}{2} N^{-2} \{\partial_0 g_{ij} - g_{hj} \partial_i \beta^h - g_{ih} \partial_j \beta^h) \tag{3.9}$$

Using the expression (3.4) for ∂_0, we obtain that

$$\omega^0_{ij} \equiv \frac{1}{2} N^{-2} \bar{\partial}_0 g_{ij}, \tag{3.10}$$

where the operator $\bar{\partial}_0$ is defined on any t-dependent space tensor $\mathbf{T}$ by the formula

$$\bar{\partial}_0 := \frac{\partial}{\partial t} - \bar{L}_\beta, \tag{3.11}$$

where $\bar{L}_\beta$ is the Lie derivative on M_t with respect to the spatial vector β. Note that $\bar{\partial}_0 \mathbf{T}$ is a t-dependent space tensor of the same type as $\mathbf{T}$. One sets

$$K_{ij} \equiv -\omega^0_{ij} n_0 \equiv -N \omega^0_{ij} \equiv -\frac{1}{2} N^{-1} \bar{\partial}_0 g_{ij} \tag{3.12}$$

and we call the t-dependent space tensor K the **extrinsic curvature** of M_t.

The remaining connection coefficients are found to be (indices raised with $\bar{g}$)

$$\omega^i_{0j} \equiv -N K^i_j + \partial_j \beta^i, \quad \omega^i_{j0} \equiv -N K^i_j. \tag{3.13}$$

The trace in the metric $\bar{g}$ of the extrinsic curvature, often denoted τ, is called the **mean curvature** of M_t:

$$\tau \equiv tr_{\bar{g}} K \equiv \bar{g}^{ij} K_{ij}. \tag{3.14}$$

3.4 *Extrinsic curvature*

We elaborate on the extrinsic curvature K (also called the **second fundamental form**) of M_t which plays an important role in the initial value formulation and other geometric problems of General Relativity,

Lemma 3.1 *The extrinsic curvature K of M_t is a covariant, symmetric 2-tensor equal to the projection on M_t of the spacetime gradient of the past oriented*[4] *unit normal n to M_t.*

The mean extrinsic curvature $\tau := Tr_{\bar{g}} K$ of K on M_t is equal to the spacetime divergence of the past directed unit normal to M_t.

Proof In a Cauchy adapted frame it holds that $n_i = n^i = 0$, $n_0 = N$. The projection on M_t of the tensor $\nabla_\alpha n_\beta$ is

$$\nabla_i n_j = -\omega^0_{ij} n_0 = -\omega^0_{ij} N \equiv K_{ij}. \tag{3.15}$$

Since $n^\lambda n_\lambda = -1$, it holds that $n_\lambda \nabla_\alpha n^\lambda = \nabla_\alpha n^0 = 0$; hence[5]

$$\nabla_\alpha n^\alpha = \nabla_i n^i = \bar{g}^{ij} K_{ij} := \tau. \tag{3.16}$$

□

Remark 3.2 *The tensor K is symmetric. It can equivalently be defined, up to a factor $1/2$, as the Lie derivative of the spacetime metric in the direction of n. It does not depend on the value of n outside of M_t.*

Remark 3.3 *A positive value of τ signals a positive divergence of the past-directed normals. A negative value of τ corresponds to the convergence of the past-directed normals, hence future expansion of the submanifolds M_t.*

3.5 *Splitting of the Riemann tensor*

We deduce from the general formula giving the Riemann tensor and the splitting of the connection the following identities[6]:

$$R_{ij,kl} \equiv \bar{R}_{ij,kl} + K_{ik}K_{lj} - K_{il}K_{kj}, \tag{3.17}$$

$$R_{0i,jk} \equiv N(\bar{\nabla}_j K_{ki} - \bar{\nabla}_k K_{ji}), \tag{3.18}$$

$$R_{0i,0j} \equiv N(\bar{\partial}_0 K_{ij} + N K_{ik} K^k{}_j + \bar{\nabla}_i \partial_j N) \tag{3.19}$$

[4] Some authors take the opposite orientation of n to define K. We follow Choquet-Bruhat, Y. and York, J. W. (1980) The Cauchy problem, in *General Relativity and Gravitation* (ed. A. Held), Plenum and Wald, R. (1984) *General Relativity*, University of Chicago Press.

[5] Recall that in a Cauchy adapted frame $\bar{g}^{ij} = g^{ij}$.

[6] Fourès (Choquet)-Bruhat, Y. (1956) *J. Rat. Mech. Anal.*, **5**, 951–66.

From these formulae one obtains the following expressions for the Ricci curvature

$$NR_{ij} \equiv N\bar{R}_{ij} - \bar{\partial}_0 K_{ij} + NK_{ij}K^h_h - 2NK_{ik}K^k{}_j - \bar{\nabla}_i \partial_j N, \tag{3.20}$$

$$N^{-1}R_{0j} \equiv \partial_j K^h_h - \bar{\nabla}_h K^h{}_j, \tag{3.21}$$

$$R_{00} \equiv N(\bar{\partial}_0 K^h_h - NK_{ij}K^{ij} + \bar{\Delta}N). \tag{3.22}$$

Also, with $\bar{R} := g^{ij}\bar{R}_{ij}$,

$$g^{ij}R_{ij} = \bar{R} - N^{-1}\bar{\partial}_0 K^h_h + (K^h_h)^2 - N^{-1}\bar{\Delta}N, \tag{3.23}$$

$$R \equiv -N^{-2}R_{00} + g^{ij}R_{ij} = \bar{R} + K_{ij}K^{ij} + (K^h_h)^2 - 2N^{-1}\bar{\partial}_0 K^h_h - 2N^{-1}\bar{\Delta}N, \tag{3.24}$$

and

$$S_{00} \equiv R_{00} - \frac{1}{2}g_{00}R \equiv \frac{1}{2}(R_{00} + g^{ij}R_{ij}).$$

It follows that

$$2N^{-2}S_{00} \equiv -2S^0_0 \equiv \bar{R} - K_{ij}K^{ij} + (K^h_h)^2. \tag{3.25}$$

Remark 3.4 *A variant of Equations (3.20) is the following (indices raised with $g^{ij} = \bar{g}^{ij}$ in a Cauchy adapted frame)*

$$NR^j_i = N\bar{R}^j_i - \bar{\partial}_0 K^j_i + NK^j_i K^h_h - \bar{\nabla}_i \partial^j N, \tag{3.26}$$

which are deduced from (3.20) and from the relation, consequent of (3.14),

$$\bar{\partial}_0 g^{ij} = 2NK^{ij} \tag{3.27}$$

4 Constraints and evolution

4.1 *Equations. Conservation of constraints*

We see in the above decomposition of the Ricci tensor that none of the components of the Einstein tensor contains the time derivatives of the lapse N or the shift β. One is thus led to consider the Einstein equations as a dynamical system for the two fundamental tensors $\bar{g}$ and K of the space slices M_t. This dynamical system splits as follows.

Constraints

The restriction to M_t of the right-hand side of the identities (3.21) and (3.25) contains only the metric $\bar{g}$ and the extrinsic curvature K of M_t as tensor fields on M_t. When the Einstein equations are satisfied, i.e. when

$$S_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\ R = T_{\alpha\beta} \equiv \rho_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\rho, \tag{4.1}$$

these identities lead to the following equations called constraints.

Momentum constraint

$$C_i \equiv \frac{1}{N}(R_{0i} - \rho_{0i}) \equiv -\bar{\nabla}_h K_i^h + \bar{\nabla}_i K_h^h - N^{-1}\rho_{0i} = 0. \tag{4.2}$$

Hamiltonian constraint

$$C_0 \equiv \frac{2}{N^2}(S_{00} - T_{00}) \equiv \bar{R} - K_j^i K_i^j + (K_h^h)^2 + 2T_0^0 = 0. \tag{4.3}$$

These constraints are transformed into a system of elliptic equations on each submanifold M_t, in particular on M_0 for $\bar{g} = g_0$, $K = K_0$, by the conformal method (see Chapter 7).

Evolution

The equations

$$R_{ij} \equiv \bar{R}_{ij} - \frac{\bar{\partial}_0 K_{ij}}{N} - 2K_{jh}K_i^h + K_{ij}K_h^h - \frac{\bar{\nabla}_j \partial_i N}{N} = \rho_{ij} \tag{4.4}$$

together with the definition

$$\bar{\partial}_0 g_{ij} = -2NK_{ij}$$

determine the derivatives transversal to M_t of $\bar{g}$ and K when these tensors are known on M_t as well as the lapse N and shift β, and the source ρ_{ij}. It is natural to look at these equations as evolution equations determining $\bar{g}$ and K, while N and β, which are projections of the tangent vector to the time line respectively on e_0 and on the tangent space to M, are considered as gauge variables. This point of view is supported by the following theorem (Anderson and York, 1997, previously given for equations with sources[7] in Choquet-Bruhat and Noutchegueme, 1986).

Theorem 4.1 *If $R_{ij} - \rho_{ij} = 0$ holds, then the constraints satisfy a linear homogeneous first order symmetric hyperbolic system*[8]*. If satisfied initially, then, they are satisfied for all time.*

Proof If $R_{ij} - \rho_{ij} = 0$ we have, in the Cauchy adapted frame, with $\rho := g^{\alpha\beta}\rho_{\alpha\beta}$,

$$R - \rho = -N^2(R^{00} - \rho^{00}).$$

Hence

$$S^{00} - T^{00} = \frac{1}{2}(R^{00} - \rho^{00}) \quad \text{and} \quad R - \rho = -2N^2(S^{00} - T^{00}) = 2(S_0^0 - T_0^0)$$

and

$$S^{ij} - T^{ij} = -\frac{1}{2}\bar{g}^{ij}(R - \rho) = -\bar{g}^{ij}(S_0^0 - T_0^0).$$

[7] Choquet-Bruhat, Y. and Noutchegueme, N. (1986) *C. R. Acad. Sci. Paris*, **303**, serie I 6, 259–63.

[8] See Section IV.4 in Appendix IV.

With these identities we may derive from the Bianchi identities a linear homogeneous system for $\Sigma_0^i \equiv S_0^i - T_0^i$ and for $\Sigma_0^0 \equiv S_0^0 - T_0^0$ with principal parts

$$N^{-2}\partial_0\Sigma_0^i + \bar{g}^{ij}\partial_j\Sigma_0^0, \quad \text{and} \quad \partial_0\Sigma_0^0 + \partial_i\Sigma_0^i.$$

Since this system is symmetrizable hyperbolic, it has a unique solution which is zero if the initial values are zero. The characteristic which determines the domain of dependence is the light cone. □

5 Hamiltonian and symplectic formulation

5.1 *Hamilton equations*

In the vacuum case the $3+1$ splitting has been given an interesting dynamical interpretation through the introduction of a Hamiltonian functional by Arnowitt *et al.* (1962)[9], following work by Dirac (1959)[10]. This dynamical system has been cast in symplectic form by Fisher and Marsden (1975)[11], using the results of Moncrief (1975)[12]. A new formulation, inspired by the $3+1$ splitting of the previous section was given in 1999 by Anderson and York[13]. A Hamiltonian formulation holds in arbitrary dimension, as we now show.

The idea is to write the Lagrangian of vacuum General Relativity,

$$L(g) := \int R\mu_g, \tag{5.1}$$

in typical Hamiltonian form.

For finite-dimensional systems, a Hamiltonian appears when the Lagrangian reads

$$\int_{t_1}^{t_2} \{p\dot{q} - H(p,q)\}dt \tag{5.2}$$

where the q are the state variables, $\dot{q}$ are their time derivatives, and p are their conjugate momenta, such that the Euler–Lagrange dynamical equations, for the q taking specified values at t_1 and t_2, can be written in the form

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{5.3}$$

In General Relativity the problem is complicated, both by the infinite dimension of the space of state variables, and by the invariance under diffeomorphisms and the existence of constraints.

[9] Arnowitt, R., Deser, S., and Misner, C. W. (1962) In *Gravitation, an Introduction to Current Research* (ed. L. Witten), Wiley.
[10] Dirac, P. A. M. (1959) *Phys. Rev.*, **114**, 924–30.
[11] Fisher, A. and Marsden, J. (1975) *Proc. Symp. Pure Math. AMS*, **27**, part 2, 219–63.
[12] Moncrief, V. (1975) *J. Math. Phys.*, **16**, 493–8 and **17**, 1893–902.
[13] Anderson, A. and York, J. (1999) *Phys. Rev. Lett.*, **82**, 4384–7.

The $n+1$ splitting leads to the choice of $\bar{g}$ and K as dynamical variables, as discussed above. The identity (3.24) giving an expression of R, and the relation $\mu_g \equiv N\mu_{\bar{g}}$ between the volume elements of g and $\bar{g}$, show that the Lagrangian of vacuum GR on a sliced manifold $(M \times R, g)$ reads, with $\tau := g^{ij}K_{ij} \equiv K^h_h$:

$$\begin{aligned} L(g) &:= \int_{t_1}^{t_2}\int_M RN\mu_{\bar{g}}dt \qquad (5.4)\\ &\equiv \int_{t_1}^{t_2}\int_M \{N(\bar{R}+|K|^2+\tau^2) - 2\partial_0\tau - 2\bar{\Delta}N\}\mu_{\bar{g}}dt. \end{aligned}$$

Using the definition (recall that $\mathcal{H}=0$ is the Hamiltonian constraint)

$$\mathcal{H} := \bar{R} - |K|^2 + \tau^2, \qquad (5.5)$$

$L(g)$ can be written as

$$L(g) = \int_{t_1}^{t_2}\int_M \{2N|K|^2 + N\mathcal{H} - 2\partial_0\tau - 2\bar{\Delta}N\}\mu_{\bar{g}}dt. \qquad (5.6)$$

By the definition of $\mu_{\bar{g}}$ and ∂_0, it holds that

$$\int_{t_1}^{t_2}\int_M \partial_0\tau\mu_{\bar{g}}dt \equiv \int_{t_1}^{t_2}\int_M \left\{\frac{\partial\tau}{\partial t} - \beta^i\frac{\partial\tau}{\partial x^i}\right\}(det\bar{g})^{\frac{1}{2}}dx^1\cdots dx^n dt. \qquad (5.7)$$

Then by the Leibnitz formula, this integral is identical with the following:

$$\begin{aligned} \int_{t_1}^{t_2}\int_M \Big\{\frac{\partial}{\partial t}[\tau(det\bar{g})^{\frac{1}{2}}] + \Big[-\tau(det\bar{g})^{-\frac{1}{2}}\frac{\partial}{\partial t}(det\bar{g})^{\frac{1}{2}} \qquad (5.8)\\ - \bar{\nabla}_i(\beta^i\tau) + \tau\bar{\nabla}_i\beta^i\Big](det\bar{g})^{\frac{1}{2}}\Big\}dx^1\cdots dx^n dt. \end{aligned}$$

The formula for the derivative of a determinant and the definition of $\bar{\partial}_0$ give that

$$(det\bar{g})^{-\frac{1}{2}}\frac{\partial}{\partial t}(det\bar{g})^{\frac{1}{2}} - \bar{\nabla}_i\beta^i \equiv \frac{1}{2}g^{ij}\bar{\partial}_0 g_{ij}, \qquad (5.9)$$

We discard in the integral (5.8) both a derivative with respect to t which vanishes for variations of fields with specified values on M_{t_1} and M_{t_2}, and a space divergence whose integral vanishes for variations with compact support in space. We denote by $\cong$ equalities obtained when such terms are discarded. We find that

$$\int \partial_0\tau\mu_{\bar{g}}dt \;\cong\; -\frac{1}{2}\int \tau g^{ij}\bar{\partial}_0 g_{ij}\mu_{\bar{g}}dt. \qquad (5.10)$$

Inserting this value in (5.6) gives

$$L(g) \;\cong\; \int_{t_1}^{t_2}\int_M \{2NK^{ij}K_{ij} + N\mathcal{H} + \tau g^{ij}\bar{\partial}_0 g_{ij}\}\mu_{\bar{g}}dt. \qquad (5.11)$$

We use the definition $-2NK_{ij} = \bar{\partial}_0 g_{ij}$, and we set

$$P^{ij} := -K^{ij} + g^{ij}\tau,$$

hence

$$P_h^h \equiv (n-1)\tau, \quad K_{ij} \equiv -P_{ij} + \frac{1}{n-1} g_{ij} P_h^h \tag{5.12}$$

We then find that the Lagrangian of vacuum General Relativity reads[14]:

$$\begin{aligned} L(g) &= \int_{t_1}^{t_2} \int_M \{P^{ij}\bar{\partial}_0 g_{ij} + N\mathcal{H}\}\mu_{\bar{g}} dt \\ &\equiv \int_{t_1}^{t_2} \int_M \left\{ P^{ij} \left(\frac{\partial}{\partial t} g_{ij} - 2\bar{\nabla}_j \beta_i \right) + N\mathcal{H} \right\} \mu_{\bar{g}} dt. \end{aligned} \tag{5.13}$$

Using again integration by parts we obtain:

$$L(g) \doteq \int \left\{ P^{ij} \frac{\partial}{\partial t} g_{ij} - 2\beta_i \mathcal{M}^i + N\mathcal{H} \right\} \mu_{\bar{g}} dt, \tag{5.14}$$

because the left-hand side of the momentum constraint is, in our notations:

$$\mathcal{M}^j := -\bar{\nabla}_i P^{ij}. \tag{5.15}$$

The Lagrangian takes for the dynamical variables $\bar{g}$ and P a form familiar in Hamiltonian theories. If we discard the integrals on $(M, \bar{g})$ of divergences of vector fields, which vanish for variations with compact support[15], it reads as the following functional of $\bar{g}$, P, N, β:

$$L(\bar{g}, P, N, \beta) := \int_{t_1}^{t_2} \int_M \left\{ P^{ij} \frac{\partial}{\partial t} g_{ij} - 2\beta_i \mathcal{M}^i + N\mathcal{H} \right\} d\mu_{\bar{g}} dt \tag{5.16}$$

corresponding to the Hamiltonian functional

$$\mathcal{F} := \int_M \mathcal{H}_{tot} \mu_{\bar{g}} \quad \text{with} \quad \mathcal{H}_{tot} := 2\beta_i \mathcal{M}^i - N\mathcal{H}. \tag{5.17}$$

In terms of P, the left-hand side $\mathcal{H}$ of the Hamiltonian constraint is calculated to be

$$\mathcal{H} \equiv \bar{R} - P^{ij} P_{ij} + \frac{1}{n-1} (P_h^h)^2. \tag{5.18}$$

The definition of the functional partial derivatives of L with respect to $\bar{g}$, P, β, and N is that the first variation of L (linearization) is given by linear maps on the first variations $\delta\bar{g}$, $\delta P, \delta\beta$, N:

$$\delta L \equiv \int_M \left\{ \frac{\partial L}{\partial g_{ij}} \delta g_{ij} + \frac{\partial L}{\partial P^{ij}} \delta P^{ij} + \frac{\partial L}{\partial \beta^i} \delta\beta_i + \frac{\partial L}{\partial N} \delta\mathcal{N} \right\} \mu_{\bar{g}} dt. \tag{5.19}$$

[14] For fields with given values on M_{t_1} and M_{t_2}.

[15] When integrals of divergences on M are not discarded one adds to the Hamiltonian an $n-1$ integral which is used to define energy and momentum in non-compact spaces.

The equation $\delta_P L = 0$ reads

$$\delta_P L \equiv \int_M \left\{ \frac{\partial g_{ij}}{\partial t} - \frac{\partial \mathcal{H}_{tot}}{\partial P^{ij}} \right\} \delta P^{ij} \mu_{\bar{g}} = 0, \tag{5.20}$$

equivalently

$$\frac{\partial g_{ij}}{\partial t} = \frac{\partial \mathcal{H}_{tot}}{\partial P^{ij}}. \tag{5.21}$$

It holds that

$$\int_M \frac{\partial \mathcal{H}_{tot}}{\partial P^{ij}} \delta P^{ij} \mu_{\bar{g}} = \int_M \{-2\beta_i \bar{\nabla}_j \delta P^{ij} - N \delta_P \mathcal{H}\} \mu_{\bar{g}}, \tag{5.22}$$

with

$$\delta_P \mathcal{H} \equiv -2 \left(P_{ij} - \frac{1}{n-1} g_{ij} P_h^h \right) \delta P^{ij} \equiv 2 K_{ij} \delta P^{ij}. \tag{5.23}$$

hence, after an integration by parts, for variations with compact support

$$\frac{\partial \mathcal{H}_{tot}}{\partial P^{ij}} = \bar{\nabla}_j \beta_i + \bar{\nabla}_i \beta_j - 2NK_{ij}. \tag{5.24}$$

This formula shows that the equality (5.21) is identical to the definition of K,

$$\bar{\partial}_0 g_{ij} = -2NK_{ij}. \tag{5.25}$$

The critical points of L considered as a functional of $\bar{g}, P, N, \beta$, or of $g = (\bar{g}, N, \beta)$ are therefore identical. The equations they satisfy are (5.21) and

$$\frac{\partial P^{ij}}{\partial t} = -\frac{\partial \mathcal{H}_{tot}}{\partial g_{ij}}, \tag{5.26}$$

together with $\delta_\beta L = 0$, $\delta_N L = 0$. These last equations give respectively the vanishing of the momentum and Hamiltonian constraints. Thus β and N appear as Lagrange multipliers for a constrained Lagrangian.

5.2 *Symplectic formulation*

By the previous definitions it holds that

$$\begin{aligned} \delta_{\bar{g},P} \mathcal{F} &\equiv \int_M \left(\frac{\partial \mathcal{H}_{tot}}{\partial g_{ij}} \delta g_{ij} + \frac{\partial \mathcal{H}_{tot}}{\partial P^{ij}} \delta P^{ij} \right) d\mu_{\bar{g}} dt \\ &\equiv \int_M (2\beta_i \delta_{\bar{g},P} \mathcal{M}^i - N \delta_{\bar{g},P} \mathcal{H}) d\mu_{\bar{g}} dt. \end{aligned} \tag{5.27}$$

We denote by Φ the constraints map $\Phi : (\bar{g}, P) \mapsto (-\mathcal{H}, 2\mathcal{M})$, mapping a pair of symmetric 2-tensor fields into a pair composed of a scalar field and a vector field. The derivative $\Phi'_{\bar{g},P}$ at a pair $(\bar{g}, P)$ is a linear map acting on a pair $(\delta\bar{g}, \delta P)$, the result being the pair $(2\delta_{\bar{g},P}\mathcal{M}, -\delta_{\bar{g},P}\mathcal{H})$. The relation (5.27) reads

$$\delta_{\bar{g},P} \mathcal{F} \equiv \langle (\beta, N), (\Phi'.(\delta\bar{g}, \delta P)) \rangle \tag{5.28}$$

where $\langle ., . \rangle$ denotes the L^2 scalar product of spatial tensors in the metric $\bar{g}$. The adjoint of the operator Φ' is a linear operator $(\Phi')^*$ from a pair consisting of a vector field β and a scalar function N, into a pair of symmetric 2-tensors. It is such that, for any set $(\beta, N; \delta\bar{g}, \delta P)$ with compact support, it holds that

$$\delta_{\bar{g},P}\mathcal{F} \equiv \langle (\beta, N), \Phi'.(\delta\bar{g}, \delta P) \rangle \equiv \langle \Phi'^*.(\beta, N), (\delta\bar{g}, \delta P) \rangle. \tag{5.29}$$

We conclude from this identity, from (5.27), and from Hamilton's equations (5.21) and (5.26) that for all $\delta\bar{g}, \delta P$ it holds that:

$$\langle \Phi'^*.(\beta, N), (\delta\bar{g}, \delta P) \rangle = \left\langle \left(-\frac{\partial P}{\partial t}, \frac{\partial \bar{g}}{\partial t} \right), (\delta\bar{g}, \delta P) \right\rangle. \tag{5.30}$$

Therefore Hamilton's equations can be written as

$$\left(\frac{\partial \bar{g}}{\partial t}, \frac{\partial P}{\partial t} \right) = J \ \Phi'^*.(\beta, N), \tag{5.31}$$

where J is the symplectic matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.32}$$

This property of the constraint map is used in the study of the linearization stability of the constraints (Chapter 7).

Remark 5.1 *The Lagrangian formulation of Einstein equations is invariant under spacetime diffeomorphisms. It is not the case for Hamiltonian formulations which suppose a slicing of the spacetime. The dynamical variables $\bar{g}$ and K (equivalently $\bar{g}$ and P) associated to a slicing have an intrinsic geometric meaning, while the other quantities used to complete the spacetime metric depend on the choice of time lines. When the slicing is chosen, the constraints are geometrically defined, but the evolution equations can take different forms, in particular by combination with the constraints. The traditional choice of lapse N and shift β, as state variables independent of $\bar{g}$, has the advantage that the right-hand side of the "evolution equations" now includes the adjoint operator of the differential of the constraints map. However, they have the disadvantage of not being equivalent to the equations $R_{ij} = 0$, which have a comparatively simple expression, and whose verification implies that the constraints satisfy a symmetric hyperbolic system. Anderson and York (see also Teitelboim and Ashtekar) have shown that by replacing N by the densitized lapse $\alpha := N(det\bar{g})^{-\frac{1}{2}}$ (which also appears in hyperbolic formulations), the evolution equations become equivalent to the equations $R_{ij} = 0$.*

6 Cauchy problem

6.1 *Definitions*

Since General Relativity is a geometric theory, problems and results from the theory should be enunciated in a coordinate independent way, even if special

coordinate systems are used to carry out proofs. We will formulate the Cauchy problem for the Einstein equations intrinsically through the following definitions:

Definition 6.1

1. *An **initial data set** is a triplet $(M, \bar{g}, K)$ where M is an n-dimensional smooth manifold, $\bar{g}$ is a properly Riemannian metric on M and K a symmetric 2-tensor.*
2. *A **development** of an initial data set is a spacetime (V, g), such that there exists an embedding i of M into V enjoying the two following properties:*
   (a) *The metric $\bar{g}$ is the pullback of g by i, $\bar{g} = i^* g$. In other words, if M is identified with its image $i(M) := M_0$ in V, $\bar{g}$ is the metric induced by g on M_0.*
   (b) *The image by i of K is the second fundamental form (extrinsic curvature) of $i(M)$ as submanifold of (V, g).*
3. *A development (V, g) of $(M, \bar{g}, K)$ is called an **Einsteinian development** if the metric g satisfies on V the Einstein equations.*

We have shown that the initial data $(\bar{g}, K)$ cannot be chosen arbitrarily; they must satisfy on M the constraints (4.2) and (4.3).

The definition of K can be written in the form

$$\partial_t g_{ij} = -2NK_{ij} + \bar{\nabla}_i \beta_j + \bar{\nabla}_j \beta_i. \tag{6.1}$$

The evolution equations (4.4) read

$$\begin{aligned} \partial_t K_{ij} = N\{\bar{R}_{ij} - 2K_{ih}K^h_j + K_{ij}K^h_h\} - \bar{\nabla}_i \partial_j N \\ + \beta^h \bar{\nabla} K_{ij} + K_{ih}\bar{\nabla}_j \beta^h + K_{hj}\bar{\nabla}_i \beta^h - N\rho_{ij}. \end{aligned} \tag{6.2}$$

The system (6.1), (6.2) is usually called the **ADM** equations. None of these equations contains the time derivatives of the lapse N or the shift β.

6.2 *The analytic case*

Lemma 6.2 *If the lapse N and the shift β are specified on V, analytic (this is the ADM gauge), then the system (6.1), (6.2) is a Cauchy–Kovalevski type system for $\bar{g}$ and K, as time-dependent tensors on M.*

Proof Definition of a Cauchy–Kovalevski system: the time derivative of each unknown is equal to an analytic function of the coordinates, the set of unknowns and their space derivatives (Appendix V). □

From this lemma, and from the Cauchy–Kowalewski theorem, one obtains[16].

[16] Darmois, G. (1923) *Les équations de la gravitation Einsteinienne*, Gauthier-Villars (case $N = 1$, $\beta = 0$). Lichnerowicz, A. (1939) *Problèmes Globaux en Mécanique Relativiste*, Herman (case $\beta = 0$), N arbitrary.

Theorem 6.3 *If the initial data are analytic on M_0 and if the sources, the shift and the lapse are analytic in a neighbourhood of M_0 in $M \times R$, then there exists a neighbourhood of M_0 in $M \times R$ such that the evolution equations (6.1), (6.2) have an analytic solution in this neighbourhood consistent with these Cauchy data.*

We deduce from this theorem and Theorem 4.1 the following one

Theorem 6.4 *Analytic initial data admits a vacuum Einsteinian development if and only if these initial date satisfy the vacuum constraints.*

Proof The "only if" part is a trivial consequence of the splitting of the Einstein tensor.

To prove the "if" part, take an arbitrary analytic shift β and an analytic lapse $N > 0$. The analytic solution of the evolution taking the specified Cauchy data satisfies the constraints in a neighbourhood of the initial surface M_0; hence the full Einstein equations in this neighbourhood. □

Remark 6.5 *A choice of lapse and shift is a gauge choice, that is a choice of a particular representative of geometric objects, here the spacetime metric.*

In the presence of sources the Einstein equations are coupled with equations satisfied by these sources, which imply the conservation laws (2.24). If the coupled system of evolution equations (6.1), (6.2) and equations satisfied by the sources is of the Cauchy–Kovalevski type, then Theorem 6.3 applies to equations with sources, and initial data satisfying the corresponding constraints.

We will treat later the local and global geometric uniqueness problems, in the more general context of non-analytic initial data.

7 Wave gauges

A solution of the Einstein equations should exhibit causal propagation. The spacetime metric at a point x should depend only on the past of this point; that is on the initial values at points which can reach x by future directed timelike or null curves. Since analytic functions are completely determined by their values in an open set, they are not well adapted to causal theories. We must look for existence and uniqueness of Einsteinian developments of non-analytic initial data sets. Due to invariance by isometries, we know that a geometric Einsteinian development admits many representatives. In particular it is represented by different sets of functions in different local coordinates. A representative in a given class is obtained by what is called a gauge choice, which gives a particular form to the equations. Since it is known that hyperbolic differential systems exhibit causal propagation, one would like a gauge choice to lead to a hyperbolic system for the Einstein equations in the chosen gauge. The data of arbitrary lapse and shift is a gauge choice, which is not geometrically linked with the $n+1$ spacetime structure, and leads to some difficulties to which we return in Chapter 8. On the other hand another gauge choice, linked with the wave equation associated to the spacetime, has been known for a long time to lead to satisfactory local existence

results[17]. The recent improvements obtained by Klainerman and Rodnianski[18] (2002) concerning the regularity of the Cauchy data, rely on this harmonic gauge, which they justly called a "wave gauge". Remarkably, the non-linear global stability of Minkowski spacetime, originally proved by Christodoulou and Klainerman (1993)[19] through equations satisfied by the Riemann tensor has now been proved by Lindblad and Rodnianski[20] by using the wave gauge.

Numerical computation[21] is now using successfully the Einstein equations in generalized wave gauge.

7.1 *Wave coordinates*

A local coordinate x^α is said to be a **wave (or harmonic) coordinate** in the open set U if and only if the scalar function $u : U \to R$ by $x \mapsto x^\alpha$ is a harmonic function for the metric g, i.e. satisfies the wave equation:

$$g^{\lambda\mu}\nabla_\lambda\partial_\mu u \equiv g^{\lambda\mu}(\partial^2_{\lambda\mu}u - \Gamma^\rho_{\lambda\mu}\partial_\rho u) \equiv -g^{\lambda\mu}\Gamma^\alpha_{\lambda\mu} = 0. \tag{7.1}$$

Indeed, for this function u we have $\partial_\rho u = \delta^\alpha_\rho$, therefore the local coordinate x^α is harmonic in U for the spacetime metric g if and only if the Christoffel symbols of this metric satisfy the following equation in U

$$F^\alpha \equiv g^{\lambda\mu}\Gamma^\alpha_{\lambda\mu} = 0. \tag{7.2}$$

We show that in wave coordinates the only second-order derivatives of the components of the metric g which appear in the component $R_{\alpha\beta}$ are derivatives of $g_{\alpha\beta}$.

Theorem 7.1 *If the $n+1$ local coordinates are harmonic for the metric g the Ricci tensor of this metric reads in these coordinates as a quasidiagonal*[22] *second-order operator for the components of g; that is, the following identity holds:*

$$R^{(h)}_{\alpha\beta} \equiv -\frac{1}{2}g^{\lambda\mu}\partial^2_{\lambda\mu}g_{\alpha\beta} + P_{\alpha\beta}(g, \partial g) \tag{7.3}$$

with P a homogeneous quadratic form in the first partial derivatives of the components of the metric g.

[17] Fourès (Choquet)-Bruhat, Y. (1952) *Acta Mathematica*, **88**, 141–225.

[18] Klainerman, S. and Rodnianski, I. (2002) *C. R. Acad. Sci. Paris*, Serie I, **334**, 125–30; see also arXiv:math. AP/109173 and 74.

[19] Christodoulou, D. and Klainerman, S. (1993) *The Global Non-Linear Stability of Minkowski Space*, Princeton Mathematical series 41.

[20] Lindblad, H. and Rodnianski, I. (2005) *Commun. Math. Phys.*, **256**(1), 43–110.

[21] Pretorius, F. (2005) *Class. Quant. Grav.*, **22** 425–52; Lindblom, L., Matthews, K., Rinne, O., and Scheel, M. (2008) gr-qc 0711-2084.

[22] That is, the principal, here second-order, terms, form a diagonal matrix (see Appendices III and IV).

Proof A straightforward computation using the expressions for the Ricci tensor and the Christoffel symbols gives the identity

$$R_{\alpha\beta} \equiv R^{(h)}_{\alpha\beta} + L_{\alpha\beta} \tag{7.4}$$

with

$$L_{\alpha\beta} \equiv \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta F^\lambda + g_{\beta\lambda}\partial_\alpha F^\lambda\} \tag{7.5}$$

with F^α given by (7.2) and

$$R^{(h)}_{\alpha\beta} \equiv -\frac{1}{2}g^{\lambda\mu}\partial^2_{\lambda\mu}g_{\alpha\beta} + P^{\rho\sigma\gamma\delta\lambda\mu}_{\alpha\beta}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu}, \tag{7.6}$$

where P is a polynomial in g and its contravariant associate, deduced from the identity

$$P^{\rho\sigma\gamma\delta\lambda\mu}_{\alpha\beta}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu} \equiv -\frac{1}{2}(\partial_\beta g^{\lambda\mu}\partial_\lambda g_{\alpha\mu} + \partial_\alpha g^{\lambda\mu}\partial_\lambda g_{\beta\mu}) - \Gamma^\mu_{\alpha\lambda}\Gamma^\lambda_{\beta\mu}. \tag{7.7}$$

□

Corollary 7.2 *The contravariant form of the Ricci tensor admits the decomposition*

$$R^{\alpha\beta} \equiv R^{\alpha\beta}_{(h)} + M^{\alpha\beta} \tag{7.8}$$

with

$$M^{\alpha\beta} \equiv \frac{1}{2}\{g^{\alpha\lambda}\partial_\lambda F^\beta + g^{\beta\lambda}\partial_\lambda F^\alpha\} \tag{7.9}$$

and

$$R^{\alpha\beta}_{(h)} \equiv \frac{1}{2}g^{\lambda\mu}\partial^2_{\lambda\mu}g^{\alpha\beta} + Q^{\rho\sigma\gamma\delta\lambda\mu}_{\alpha\beta}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu}. \tag{7.10}$$

Proof The existence of the decomposition (7.8) results from the identity

$$\partial_\lambda g^{\alpha\beta} \equiv -g^{\alpha\sigma}g^{\beta\rho}\partial_\lambda g_{\sigma\rho}.$$

The expression for Q can be most easily calculated by using the following form of the harmonicity conditions:

$$F^\alpha \equiv -\nabla_\lambda g^{\lambda(\alpha)} = 0 \tag{7.11}$$

where the notation (α) means that α is a numb index, not to be considered in the covariant derivation. The Ricci identity applied to the vector $g^{\lambda(\beta)}$ gives

$$\nabla_\lambda\nabla^\alpha g^{\lambda(\beta)} - \nabla^\alpha\nabla_\lambda g^{\lambda(\beta)} \equiv R_\lambda{}^{\alpha,\lambda}{}_\mu g^{\mu(\beta)} \equiv R^{\alpha\beta}. \tag{7.12}$$

That is, using symmetrization between α and β

$$R^{\alpha\beta} \equiv R^{\alpha\beta}_{(h)} + M^{\alpha\beta} \tag{7.13}$$

with

$$R^{\alpha\beta}_{(h)} \equiv \frac{1}{2}\{\nabla_\lambda \nabla^\alpha g^{\lambda(\beta)} + \nabla_\lambda \nabla^\beta g^{\lambda(\alpha)}\} \tag{7.14}$$

and (remember that the F^α are a collection of scalar functions, not the components of a vector)

$$M^{\alpha\beta} \equiv \frac{1}{2}\{g^{\alpha\lambda}\partial_\lambda F^\beta + g^{\beta\lambda}\partial_\lambda F^\alpha\}. \tag{7.15}$$

To compute $R^{\alpha\beta}_{(h)}$ we first remark that

$$\nabla^\alpha g^{\lambda(\beta)} = -g^{\alpha\rho} g^{\lambda\mu}\Gamma^{(\beta)}_{\rho\mu} = -g^{\alpha\rho} g^{\lambda\mu} g^{\gamma(\beta)}[\gamma, \rho\mu]$$

from which we deduce that

$$\begin{aligned} R^{\alpha\beta}_{(h)} &\equiv -\frac{1}{4} g^{\lambda\mu}\nabla_\lambda\{(g^{\alpha\rho} g^{(\beta)\gamma} + g^{\beta\rho} g^{(\alpha)\gamma})\partial_\mu g_{\gamma\rho}\} \\ &\equiv \frac{1}{4} g^{\lambda\mu}\nabla_\lambda\{\partial_\mu g^{\alpha(\beta)} + \partial_\mu g^{\beta(\alpha)}\}. \end{aligned}$$

Therefore, if we suppress marking numb indices, since all derivatives are now ordinary partial derivatives,we have

$$R^{\alpha\beta}_{(h)} \equiv \frac{1}{2} g^{\lambda\mu}\partial^2_{\lambda\mu} g^{\alpha\beta} + Q^{\alpha\beta,\rho\sigma\gamma\delta\lambda\mu}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu} \tag{7.16}$$

with

$$Q^{\alpha\beta,\rho\sigma\gamma\delta\lambda\mu}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu} \equiv g^{\lambda\mu}\{-2\Gamma^\rho_{\lambda\mu}\partial_\rho g^{\alpha\beta} + \Gamma^\alpha_{\lambda\rho}\partial_\mu g^{\rho\beta} + \Gamma^\beta_{\rho\lambda}\partial_\mu g^{\alpha\rho}\} \tag{7.17}$$

□

Corollary 7.3 *The Einstein tensor S satisfies the identity*

$$S^{\alpha\beta} := R^{\alpha\beta} - \frac{1}{2} R g^{\alpha\beta} \equiv S^{\alpha\beta}_{(h)} + \frac{1}{2}\{g^{\alpha\lambda}\partial_\lambda F^\beta + g^{\beta\lambda}\partial_\lambda F^\alpha - g^{\alpha\beta}\partial_\lambda F^\lambda\}. \tag{7.18}$$

where the $S^{\alpha\beta}_{(h)}$ are a quasidiagonal system of wave operators in the metric g for the contravariant tensor densities:

$$\mathcal{G}^{\alpha\beta} \equiv g^{\alpha\beta}(detg)^{1/2}. \tag{7.19}$$

Proof We have

$$S^{\alpha\beta}_{(h)} \equiv R^{\alpha\beta}_{(h)} - \frac{1}{2} g^{\alpha\beta} R_{(h)}. \tag{7.20}$$

We deduce from the formula (7.13) that

$$R \equiv R_{(h)} + L, \quad \text{with} \quad L \equiv \partial_\lambda F^\lambda, \tag{7.21}$$

and

$$R_{(h)} \equiv \frac{1}{2} g^{\lambda\mu} g_{\alpha\beta}\partial^2_{\lambda\mu} g^{\alpha\beta} + g_{\alpha\beta} Q^{\alpha\beta,\rho\sigma\gamma\delta\lambda\mu}(g)\partial_\rho g_{\gamma\delta}\partial_\sigma g_{\lambda\mu}. \tag{7.22}$$

Elementary computation gives

$$g_{\alpha\beta}\partial^2_{\lambda\mu}g^{\alpha\beta} \equiv \partial_\lambda(g_{\alpha\beta}\partial_\mu g^{\alpha\beta}) - \partial_\lambda g_{\alpha\beta}\partial_\mu g^{\alpha\beta} \tag{7.23}$$

hence, since $g_{\alpha\beta}$ and $g^{\alpha\beta}$ are inverse matrices

$$g_{\alpha\beta}\partial^2_{\lambda\mu}g^{\alpha\beta} \equiv -\partial_\lambda(g^{\alpha\beta}\partial_\mu g_{\alpha\beta}) - \partial_\lambda g_{\alpha\beta}\partial_\mu g^{\alpha\beta} \tag{7.24}$$

and, using the formula for the derivative of a determinant

$$g_{\alpha\beta}\partial^2_{\lambda\mu}g^{\alpha\beta} \equiv -\partial_\lambda(|detg|^{-1}\partial_\mu|detg|) - \partial_\lambda g_{\alpha\beta}\partial_\mu g^{\alpha\beta}. \tag{7.25}$$

We denote by $\sim$ an equality up to the addition of terms containing derivatives of order at most one. We have just proved that

$$R_{(h)} \sim -\frac{1}{2}g^{\alpha\beta}g^{\lambda\mu}\partial^2_{\lambda\mu}g_{\alpha\beta} \sim -\frac{g^{\lambda\mu}}{|\det g|^{\frac{1}{2}}}\partial^2_{\lambda\mu}|\det g|^{\frac{1}{2}}. \tag{7.26}$$

Therefore

$$S^{\alpha\beta}_{(h)} \sim \frac{1}{2}(detg)^{-1/2}g^{\lambda\mu}\partial^2_{\lambda\mu}\mathcal{G}^{\alpha\beta} \tag{7.27}$$

as claimed. □

7.2 *Generalized wave coordinates*

It has been pointed out by H. Friedrich[23] that one obtains also a quasidiagonal operator on the components of the metric for the Ricci tensor if the harmonicity conditions in the open set U are replaced by the more general ones

$$F^\alpha_H := F^\alpha - H^\alpha = 0, \quad \text{with} \quad F^\alpha \equiv g^{\lambda\mu}\Gamma^\alpha_{\lambda\mu} \tag{7.28}$$

where the H^α are given scalar functions on U. Indeed it results from (7.4) that the Ricci tensor can be written

$$R_{\alpha\beta} \equiv R^{(h)}_{\alpha\beta} + \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta(F^\lambda_H + H^\lambda) + g_{\beta\lambda}\partial_\alpha(F^\lambda_H + H^\lambda)\}. \tag{7.29}$$

If $F^\alpha_H = 0$ and if H^α is known, the above system reduces to a quasidiagonal system of quasilinear wave operators for the metric,

$$R^{(h,H)}_{\alpha\beta} \equiv R^{(h)}_{\alpha\beta} + \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta H^\lambda + g_{\beta\lambda}\partial_\alpha H^\lambda\}. \tag{7.30}$$

These operators have the same principal part as the operators $R^{(h)}_{\alpha\beta}$. Decompositions analogous to (7.13) and (7.27) hold for $R^{\alpha\beta}$ and $S^{\alpha\beta}$. Their advantage is that they contain a new set of functions H^α, which can be freely specified, or eventually chosen to satisfy *ad hoc* equations.

[23] Friedrich, H. (1985) *Commun. Math. Phys.*, **100**, 525–43.

7.3 *Damped wave coordinates*

Recently, numerical analysts[24] have been led to introduce new reduced Einstein equations in a wave-related gauge, called **damped wave coordinates**. The operator (7.3) is replaced by

$$R^{(h,H,\gamma_0)}_{\alpha\beta} \equiv R^{(h)}_{\alpha\beta} + \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta H^\lambda + g_{\beta\lambda}\partial_\alpha H^\lambda\} + \frac{1}{2}\gamma_0 F^\lambda_H (g_{\beta\lambda}n_\alpha + g_{\alpha\lambda}n_\beta - g_{\alpha\beta}n_\lambda), \tag{7.31}$$

with n the unit normal to spacelike slices and γ_0 a constant. The presence of this non-zero constant seems to stabilize the results of the numerical calculations; it determines the time rate at which the gauge conditions $F^\lambda_H = 0$ are damped under evolution.

7.4 *Globalization in space, $\hat{e}$ wave gauges*

It is well known that a change of coordinates in U can also be interpreted as a diffeomorphism from U onto another open set $\hat{U}$. We shall take here this point of view to fix on a pseudo Riemannian manifold (V, g) a gauge condition which will turn its Ricci tensor into a globally defined quasilinear, quasidiagonal second-order operator.

Let us consider on V a given metric $\hat{e}$. We say that g is in wave gauge with respect to $\hat{e}$ (cf. Hawking and Ellis, 1973) if the identity map $V \to V$ is a wave map from (V, g) into $(V, \hat{e})$. We recall that a mapping $f : V \to \hat{V}$, given in local coordinates by $(x^\alpha) \mapsto (y^A = f^A(x^\alpha))$ is called a wave map from (V, g) into $(\hat{V}, \hat{e})$ if f satisfies the semi-linear equation, with $\hat{\nabla}$ the covariant derivative of gradient of maps $(V, g) \to (\hat{V}, \hat{e})$

$$g.\hat{\nabla}^2 f = 0,$$

which in coordinate components reads as

$$g^{\alpha\beta}(\partial^2_{\alpha\beta} f^A - \Gamma^\lambda_{\alpha\beta}\partial_\lambda f^A - \Gamma^A_{BC}(f)\partial_\alpha f^B \partial_\beta f^C) = 0; \tag{7.32}$$

here $\Gamma^\lambda_{\alpha\beta}$ and Γ^A_{BC} are respectively the connections of g and $\hat{e}$.

Suppose that f is the identity map and $\hat{e}$ is a flat metric in Cartesian coordinates. Then the above equation reduces to the previously considered harmonicity condition for these coordinates. For an arbitrary metric $\hat{e}$ and for f the identity map the harmonicity condition (7.32) reads as

$$\hat{F}^\lambda \equiv g^{\alpha\beta}(\Gamma^\lambda_{\alpha\beta} - \hat{\Gamma}^\lambda_{\alpha\beta}) = 0. \tag{7.33}$$

This condition is vector-valued, coordinate-independent (the difference of two connections is a tensor), and defined on the whole of V. The Ricci tensor of a metric g in the $\hat{e}$ wave gauge is a tensor denoted $Ricc^{(\hat{e})}(g)$.

[24] Gundlach, C., Calabrese, G., Hinder, I. and Martin-Garcial, J. (2005) *Class. Quant. Grav.*, **22**, 3767–774; Pretorius, F. (2005) *Phys. Rev. Lett.*, **95**, 121; Lindblom, L., Matthews, K., Rinne, O., and Scheel, M. (2008) gr-qc 0711-2084.

If $\hat{e}$ is given the condition (7.33) on g has the same principal part as the previously considered harmonicity condition, Therefore the Ricci tensor in the $\hat{e}$ wave gauge is a quasi-linear, quasidiagonal operator on g, tensor-valued, depending on $\hat{e}$. We denote by D the Riemannian covariant derivative in the metric $\hat{e}$. We set

$$R^{(\hat{e})}_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}(g_{\alpha\lambda}D_\beta \hat{F}^\lambda + g_{\beta\lambda}D_\alpha \hat{F}^\lambda)$$

We know that

$$R^{(\hat{e})}_{\alpha\beta} \equiv -\frac{1}{2}g^{\lambda\mu}D_\lambda D_\mu g_{\alpha\beta} + f_{\alpha\beta},$$

with $f_{\alpha\beta}$ independent of the second derivatives of g.

To compute explicitly the tensor $f_{\alpha\beta}$ we choose coordinates such that at the considered point $x \in V$ the Christoffel symbols of $\hat{e}$ vanish: $\hat{\Gamma}^\lambda_{\alpha\beta} = 0$. At x we have $\partial_\alpha = D_\alpha$; therefore at x,

$$D_\lambda D_\mu g_{\alpha\beta} = \partial^2_{\lambda\mu} g_{\alpha\beta} - g_{\rho\beta}\partial_\lambda \hat{\Gamma}^\rho_{\alpha\mu} - g_{\alpha\rho}\partial_\lambda \hat{\Gamma}^\rho_{\mu\beta}. \tag{7.34}$$

Using the definition

$$\hat{F}^\beta \equiv F^\beta - g^{\lambda\mu}\hat{\Gamma}^\beta_{\lambda\mu}$$

and the expression at x for the Riemann tensor of $\hat{e}$,

$$\hat{R}_\mu{}^\lambda{}_{\alpha\beta} = \partial_\alpha \hat{\Gamma}^\lambda_{\mu\beta} - \partial_\beta \hat{\Gamma}^\lambda_{\mu\alpha},$$

together with the previously computed expression for $R_{\alpha\beta}$, we find that at an arbitrary point $f_{\alpha\beta}$ is the tensor

$$f_{\alpha\beta} \equiv P^{\rho\sigma\gamma\delta\lambda\mu}_{\alpha\beta} D_\rho g_{\gamma\delta} D_\sigma g_{\lambda\mu} + \frac{1}{2}g^{\lambda\mu}\{g_{\alpha\rho}\hat{R}_\lambda{}^\rho{}_{\beta\mu} + g_{\beta\rho}\hat{R}_\lambda{}^\rho{}_{\alpha\mu}\}.$$

The use, at x, of the equality

$$D_\lambda D_\mu g^{\alpha\beta} = \partial^2_{\lambda\mu} g^{\alpha\beta} + g^{\rho\beta}\partial_\lambda \hat{\Gamma}^\alpha_{\mu\rho} + g^{\alpha\rho}\partial_\lambda \hat{\Gamma}^\beta_{\mu\rho},$$

gives, in an analogous way, the formula (valid at any point)

$$R^{\alpha\beta} \equiv R^{\alpha\beta}_{(\hat{e})} + \frac{1}{2}\{g^{\alpha\lambda}D_\lambda \hat{F}^\beta + g^{\beta\lambda}D_\lambda \hat{F}^\alpha\} \tag{7.35}$$

where

$$R^{\alpha\beta}_{(\hat{e})} \equiv \frac{1}{2}g^{\lambda\mu}D_\lambda D_\mu g^{\alpha\beta} + Q^{\alpha\beta\rho\sigma}_{\gamma\delta\lambda\mu}(g)D_\rho g^{\gamma\delta} D_\sigma g^{\lambda\mu} - \frac{1}{2}g^{\lambda\mu}\{g^{\alpha\rho}\hat{R}^\beta_{\mu\lambda\rho} + g^{\beta\rho}\hat{R}^\alpha_{\mu\lambda\rho}\}. \tag{7.36}$$

This tensorial expression for the reduced covariant Ricci tensor is denoted in intrinsic form (a dot denotes a contracted product)

$$Ricc(g)^{(\hat{e})} \equiv -\frac{1}{2}g.D^2 g + P(g)(Dg, Dg) - g.\hat{R}iemann.$$

One can also define a generalized $\hat{e}$ wave gauge by setting $\hat{F}$ equal to a specified vector $\hat{f}$.

7.5 *Local in time existence in a wave gauge*

To formulate the intrinsic Cauchy problem for Einstein's equations in the form of standard PDE analyses we proceed as follows.

1. We embed M in a product $M \times R$. We take $t \in R$ as a time coordinate and $t = t_0$ as the equation for M. We denote by $\bar{g}_{ij}$ the components of $\bar{g}$ in local coordinates on M_{t_0}. We choose on M_{t_0} a scalar function $\bar{N} > 0$ and a tangent vector $\bar{\beta}$. With these elements we construct the restriction to M_{t_0} of a spacetime metric g, with components

$$g_{ij} = \bar{g}_{ij}, \quad g_{i0} = \bar{\beta}_i, \quad g^{00} = -\bar{N}^{-2}, \quad \text{for} \quad t = t_0,$$

such that the initial spacetime metric is of Lorentzian signature.

2. The initial values of $\partial_t g_{ij}$ are determined through their relation with K_{ij}; see equation (6.1).

3. The initial values of the derivatives of the lapse and shift are determined through the data and the initial wave gauge conditions. These conditions read in the case of the $\hat{e}$ wave gauge, with $\hat{e}$ a given smooth metric on $M \times R$ with Riemannian connection $\hat{\Gamma}$

$$\hat{F}^\alpha \equiv g^{\lambda\mu}(\Gamma^\alpha_{\lambda\mu} - \hat{\Gamma}^\alpha_{\lambda\mu}) = 0 \quad \text{for} \quad t = t_0.$$

These equations for the vector $\hat{F}$ reduce in the case $\bar{\beta} = 0$ to the following ones (cf. the expressions for the Christoffel symbols and the value for K found in previous sections):

$$-\bar{N}^{-3}(\overline{\partial_0 N} + \bar{N}\bar{g}^{ij}\bar{K}_{ij}) = g^{\lambda\mu}\hat{\Gamma}^0_{\lambda\mu}, \quad \text{for} \quad t = t_0, \tag{7.37}$$

and

$$\bar{N}^{-1}\bar{g}^{ih}\partial_h\bar{N} - \bar{N}^{-2}\bar{g}^{ih}\partial_t g_{0h} + \bar{g}^{jh}\bar{\Gamma}^i_{jh} = g^{\lambda\mu}\hat{\Gamma}^i_{\lambda\mu} \quad \text{for} \quad t = t_0. \tag{7.38}$$

In the case $\bar{\beta} \not\equiv 0$ there are analogous expressions, with additional space derivatives; they still determine the values of $\partial_t g_{00}$ and $\partial_t g_{0h}$ for $t = t_0$.

We can now use the results of Appendix III to obtain a local in time, global in space, existence theorem in the $\hat{e}$ wave gauge. We treat here the vacuum case. Einstein equations with sources are considered in subsequent chapters.

Theorem 7.4 *(Local in time existence and uniqueness in the $\hat{e}$ wave gauge). We take $\hat{e} = dt^2 + e$ with e a smooth, Sobolev regular*[25] *metric on M, used to define the Sobolev spaces*[26] *H_s and $\overset{0}{H}_s$.*

Hypothesis on the initial data set $(M, \bar{g}, K)$.

[25] See Appendix I.

[26] Recall that if M is compact or Euclidean at infinity, then $H_s \equiv \overset{0}{H}_s$.

1. *$\bar{g}$ is a continuous properly Riemannian metric on M uniformly equivalent to e and such that:*

$$D\bar{g} \in \overset{0}{H}_{s-1}, \quad \textit{with} \quad s-1 > \frac{n}{2}.$$

2. *K is a symmetric 2-tensor on M such that:*

$$K \in \overset{0}{H}_{s-1}, \quad \textit{with} \quad s-1 > \frac{n}{2}.$$

Conclusions:

The initial data set admits a development (V_T, g), $V_T \equiv M \times [0,T)$, such that the spacetime metric g is[27] *a regularly sliced Lorentzian metric in $\tilde{E}_s(T)$ and satisfies on V_T the Einstein equations in the $\hat{e}$ harmonic gauge.*

Two such developments in the $\hat{e}$ wave gauge (V_T, g_1) and (V_T, g_2), which are in $\tilde{E}_s(T)$ with $s > \frac{n}{2} + 1$, and which take the same initial values $(\bar{g}, K, \bar{N}, \bar{\beta})$ on M, coincide on V_T.

Proof The vacuum Einstein equations in a wave gauge are, for the tensor field g, a quasidiagonal, hyperquasi-linear, second-order system of the type treated in Appendix III . It satisfies the hypotheses enunciated in that appendix. We determine from the initial data $(\bar{g}, K)$ initial data for the full spacetime metric g by choosing on M_0, for instance, the lapse $\bar{N}$ to be equal to one and the shift $\bar{\beta}$ to be zero. The values $\overline{\partial_t g_{ij}}$ on M_0 of the derivatives $\partial_t g_{ij}$ are determined by the relations giving $\bar{K}$, which reduce then to

$$\overline{\partial_t g_{ij}} = -2\bar{K}_{ij} \tag{7.39}$$

We determine $\overline{\partial_t N}$ and $\overline{\partial_t \beta}$ by the initial wave gauge conditions (7.37) and (7.38).

The Sobolev multiplication properties show that the derivative of g on M_0 is then in H_{s-1} if $s > \frac{n}{2}$. The existence and uniqueness theorem for a solution of the Einstein equations reduced to a wave gauge then follows. □

We have already pointed out that the wave coordinates are a particular case of the $\hat{e}$ wave gauge. These equations as well as the reduced Einstein equations in generalized, or damped generalized, wave coordinates are all of the same type, a quasilinear, quasidiagonal system of wave equations. A similar existence and uniqueness theorem holds for them.

Remark 7.5 *We can write this existence theorem for Cauchy data in the local Sobolev spaces $H_{s,u.l} \times H_{s-1,u.l}$. We will do it for the original, geometric Einstein equations.*

[27] Recall that (see Appendix III) $\tilde{\mathrm{E}}_s(T)$ is the space of tensor fields u such that $u \in \bar{C}(V_T)$, space of continuous and bounded functions on V_T, while $\partial_t u$, $Du \in E_{s-1}(T) \equiv \bigcap_{k \leq s-1} C^k([0,T), \overset{0}{H}_{s-k-1})$.

8 Local existence for the full Einstein equations

To show that the solution obtained in a wave gauge satisfies the full Einstein equations we must show that it does satisfy the wave gauge condition. This turns out to be true if and only if the initial data satisfy the constraints.

8.1 *Preservation of the wave gauges*

We use the Bianchi identities to show that, if g is a solution of the reduced Einstein equations in a harmonic (wave) gauge, then the gauge conditions satisfy a second-order linear quasidiagonal differential system.

Recall that the Bianchi identities imply the following identities for the Einstein tensor S:

$$\nabla_\alpha S^{\alpha\beta} \equiv 0. \tag{8.1}$$

Case of wave coordinates. If g is a solution of the reduced Einstein equations

$$R^{\alpha\beta}_{(h)} = \rho^{\alpha\beta}, \tag{8.2}$$

it holds that, with $T^{\alpha\beta} := \rho^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta}\rho$,

$$S^{\alpha\beta} - T^{\alpha\beta} = -\frac{1}{2}(g^{\alpha\lambda}\partial_\lambda F^\beta + g^{\beta\lambda}\partial_\lambda F^\alpha - g^{\alpha\beta}\partial_\lambda F^\lambda) \tag{8.3}$$

where $F^\alpha := g^{\lambda\mu}\Gamma^\alpha_{\lambda\mu}$(see Equation (7.2)). An elementary computation using the Bianchi identities and the conservation laws for the source T shows then that F^α satisfies a linear system

$$g^{\alpha\lambda}\partial^2_{\alpha\lambda} F^\beta + A^{\beta\lambda}_\alpha \partial_\lambda F^\alpha = 0, \tag{8.4}$$

where the As are linear functions in the Christoffel symbols of g.

The same computation which gives Equation (8.4) shows that, if g is a solution of the Einstein equation in *generalized wave coordinates*[28]

$$R^{(h)}_{\alpha\beta} + \frac{1}{2}\{g_{\alpha\lambda}\partial_\beta H^\lambda + g_{\beta\lambda}\partial_\alpha H^\lambda\} = \rho_{\alpha\beta}, \tag{8.5}$$

it holds that:

$$g^{\alpha\lambda}\partial^2_{\alpha\lambda} F^\beta_H + A^{\beta\lambda}_\alpha \partial_\lambda F^\alpha_H = 0. \tag{8.6}$$

A similar equation holds, with a different coefficient A, in the case of *damped generalized wave coordinates.*

A solution of the reduced Einstein equations in the $\hat{e}$ *wave gauge* is such that

$$S_{\alpha\beta} - T_{\alpha\beta} = \frac{1}{2}(g_{\alpha\lambda} D_\beta \hat{F}^\lambda + g_{\beta\lambda} D_\alpha \hat{F}^\lambda - g_{\alpha\beta} D_\lambda \hat{F}^\lambda)$$

[28] See Section 7.2.

The wave gauge vector $\hat{F}$ satisfies again a quasidiagonal linear system which can be written

$$g^{\alpha\lambda}D_\alpha D_\lambda \hat{F}^\beta + \hat{A}^{\beta\lambda}_\alpha D_\lambda \hat{F}^\alpha = 0,$$

where $\hat{A}$ is now a 3-tensor, depending only on g and Dg.

In all cases the Bianchi identities impose to the set of functions F^α, F^α_H or mappings $\hat{F}$, to satisfy a linear homogeneous system of wave equations in the spacetime metric g.

The following lemma results from the equations satisfied by F and from the uniqueness theorem proved in Appendix III:

Lemma 8.1 *A wave coordinates reduced Einsteinian development of initial data on M is an Einsteinian development*[29] *if and only if the reduced initial data are such that the set of functions F^α satisfy*

$$F^\alpha|_M = 0 \quad \text{and} \quad \partial_t F^\alpha|_M = 0. \tag{8.7}$$

Corresponding properties hold[30] *for the generalized, damped, and $\hat{e}$ wave gauges.*

We now prove the equivalence of Equations (8.7) with the constraints.

Lemma 8.2 *For a solution of the wave coordinates reduced Einstein equations, with initial data satisfying $F^\alpha|_M = 0$, the conditions $\partial_t F^\alpha|_M = 0$ are satisfied if and only if the initial data satisfy the constraints.*

Proof The Einstein constraints on an initial manifold M_{t_0} read as

$$S^{\alpha 0} - T^{\alpha 0} = 0, \quad \text{for} \quad t = t_0, \tag{8.8}$$

where the $S^{\alpha 0}$ component of the Einstein tensor is given by the identity (see 7.18)

$$S^{\alpha 0} \equiv S^{\alpha 0}_{(h)} + \frac{1}{2}\{g^{\alpha\lambda}\partial_\lambda F^0 + g^{0\lambda}\partial_\lambda F^\alpha - g^{\alpha 0}\partial_\lambda F^\lambda\}. \tag{8.9}$$

Suppose that $F^\alpha|_{t=t_0} = 0$. Then also $\partial_i F^\alpha|_{t=t_0} = 0$ and the identity (8.8) implies the equality

$$S^{\alpha 0} = S^{\alpha 0}_{(h)} + \frac{1}{2}g^{00}\partial_0 F^\alpha, \tag{8.10}$$

from which the conclusion follows. □

Analogous lemmas and proofs hold for the other types of wave gauge.

[29] In appropriate function spaces.

[30] It is meaningful to speak of global wave coordinates only if M is diffeomorphic to R^n.We write up this case to avoid cumbersome notations. The steps of the proof are the same for the other types of wave gauge that we considered.

8.2 *Geometric local existence*

Theorem 8.3 *(Solution of the full Einstein equations). If the initial data set $(M, \bar{g}, K)$ satisfies the hypothesis of the existence theorem (7.4) in some wave gauge, and the constraints, then the solution obtained in that wave gauge satisfies the full geometric vacuum Einstein equations.*

Proof We have chosen the non-geometric initial data in order to satisfy the gauge condition on the initial manifold. The theorem results then from Lemmas (8.1) and (8.2). □

8.3 *Geometric uniqueness*

Theorem 8.4 *(Local in time geometrical uniqueness). Let (V_1, g_1) and (V_2, g_2) be two solutions in $\tilde{E}_s(T)$ with $s > \frac{n}{2} + 2$ of the intrinsic Cauchy problem $(M, \bar{g}, \bar{K})$ for the vacuum Einstein equations. There exists an isometry from (U_1, g_1) onto (U_2, g_2), where U_1 and U_2 are neighbourhoods of the embedding M_0 of M, respectively in V_1 and V_2.*

We will prove the theorem by considering representatives in a wave gauge.

We first prove two lemmas. in the first lemma we show that, given a Lorentzian metric in a neighbourhood of M_0 in $M \times R$, there exists a diffeomorphism from U_1 onto U_2 which reduces to the identity on M_0 and has a gradient which takes specified values on M_0. In the second lemma, we show that if a metric g_1 is a development of an initial data set $(\bar{g}, K)$ there is a wave gauge in which the Cauchy data corresponding to a metric g isometric to g_1 take specified values which depend only on the geometric data $(\bar{g}, K)$.

Lemma 8.5 *If $g \in \tilde{E}_s(V_T)$ with $s > \frac{n}{2} + 1$ there exists a wave map f, diffeomorphism from a strip $W_\tau := M \times (-\tau, \tau)$ onto a neighbourhood U of $M \times \{0\}$ in $M \times R$, such that f takes the following initial values:*

$$f(., 0) = Identity, \quad \partial_t f(., 0) = (a, b),$$

with $a > 0$ a specified scalar on M and b a specified tangent vector to M.

Proof Suppose that M is diffeomorphic to R^n, the equations to satisfy by the mapping f are then a set of $n + 1$ scalar wave equations for the $n + 1$ scalar unknown $f^{(\lambda)} \equiv f^\lambda$ on V:

$$g^{\alpha\beta} \nabla_\alpha \partial_\beta f^{(\lambda)} = 0. \tag{8.11}$$

We take in V coordinates adapted to the product structure, that is x^i coordinates on M and $t \in R$. The initial data for f then take the following values:

$$f^j(x, 0)) = x^j, \qquad f^0(x, 0) = 0.$$

We choose for values of the time derivatives

$$\partial_t f^0(., 0) = a > 0, \quad \partial_t f^i(., 0) = b^i.$$

where $a \in \tilde{H}_s$ is a scalar function bounded away from zero and $b \in \overset{0}{H}_s$ a vector field on M. The theorem on linear hyperbolic equations shows that each of the Equations (8.1) has one and only one solution such that $\partial f^\lambda \in \tilde{E}_s(T) \subset \bar{C}^0(T)$, taking the indicated initial values, if $g \in \tilde{E}_s(T)$ and $s > \frac{n}{2} + 1$. Since ∂f is uniformly an isomorphism of the tangent space to V at each point of $M \times \{0\}$ there is a strip $W_\tau \subset V_T$ such that f is a diffeomorphism from W_τ onto a neighbourhood U of $M \times \{0\}$ in $M \times R$.

When M is not diffeomorphic to R^n one constructs a unique wave map f taking initial values which enjoy the properties specified for the set of scalar functions f^α. □

Lemma 8.6 *Let f be the unique wave map corresponding to the data (8.11). It is a diffeomorphism between neighbourhoods of M_0 in V. Let $\underline{g} := f^* g$ be the metric in wave gauge pull back by f of a metric g which is the development of a specified initial data set $(\bar{g}, K)$. It is possible to choose a and b such that the initial values of the metric in wave gauge $\underline{g}(.,0)$ and $(\partial_t \underline{g})(.,0)$ take preassigned specified values, depending only on the geometric data $\bar{g}$ and $\bar{K}$.*

Proof We set $\underline{x}^\alpha = f^\alpha(x)$. We have:

$$g_{\lambda\mu} = \frac{\partial f^\alpha}{\partial x^\lambda} \frac{\partial f^\beta}{\partial x^\mu} \underline{g}_{\alpha\beta}.$$

Therefore the components of the metrics g and g_* are linked on $M \times \{0\}$ by the following relations:

$$\underline{g}_{ij} = g_{ij}, \quad \underline{g}_{i0} = a^{-1}(g_{i0} - b^j g_{ij}),$$
$$\underline{g}_{00} = a^{-2}(g_{00} - 2\underline{g}_{0i} a b^i + g_{ij} b^i b^j)$$

If $g(.,0) \in \tilde{H}_s$, $s > \frac{n}{2} + 1$, we can find $a \in \tilde{H}_s, b \in \overset{0}{H}_s$ such that, for instance, $\underline{g}_{0i} = 0$ and $\underline{g}_{00} = -1$ on M_0, i.e. $\underline{\beta} = 0$ and $\underline{N} = 1$.

To determine the initial values of $\underline{\partial_t g}_{0\alpha}$ we use the following values on M_0 of the gradient of f^{-1} :

$$\partial_i F^j(.,0) = \delta_i^j, \quad \partial_i F^0(.,0) = 0, \quad F : f^{-1}, \tag{8.12}$$

$$\partial_t F^0(.,0) = a^{-1}(.) > 0, \quad \partial_t F^i(.,0) = -a^{-1} b^i(.). \tag{8.13}$$

The transformation law of the Christoffel symbols

$$\underline{\Gamma}^0_{ij} = \frac{\partial f^0}{\partial x^\lambda} \frac{\partial F^\alpha}{\partial \underline{x}^i} \frac{\partial F^\beta}{\partial \underline{x}^j} \Gamma^\lambda_{\alpha\beta} + \frac{\partial f^0}{\partial x^\alpha} \frac{\partial}{\partial \underline{x}^i} \frac{\partial F^\alpha}{\partial \underline{x^j}}$$

implies $\underline{\Gamma}^0_{ij} = a\Gamma^0_{ij}$ on M_0, that is $K_{ij} = \underline{K}_{ij}$, as foreseen since the extrinsic curvature is a geometrical element associated with the isometric embedding of $(M, \bar{g})$ into (V, g). One deduces from $\underline{K}_{ij} = K_{ij}$ that $\partial_t \underline{g}_{ij}$ takes values on M_0

depending only on the geometrical data and the choice of a and b; it is the same for $\partial_t \underline{g}_{0\alpha}$ which are determined by the initial wave gauge conditions.

The initial data $\underline{\partial_t g}(.,0)$ belong to[31] H_{s-1} if $(\bar{g}, \bar{K}) \in \tilde{H}_s \times H_{s-1}$ while $a \in \tilde{H}_s$, $s > \frac{n}{2} + 1$. The initial data for the mapping f are such that $f(.,0) - Id = 0$, while $\partial_t f(.,0)$ is in $\tilde{H}_s$. However, since the coefficient of ∂f in the wave, or wave map equation is only in $\overset{0}{E}_{s-1}$ if $g \in \tilde{E}_s$. The solution f of the wave, or wave map, equation is only such that $\partial f \in \tilde{E}_{s-1}$. The image $\underline{g} := f^* g$ is also only in $\tilde{E}_{s-1}$ on W_τ, though its trace on M_0 is in $\tilde{H}_s$. □

Proof of the theorem. Using the previous lemmas we construct two wave maps f_1 and f_2 such that $f_1^* g_1$ and $f_2^* g_2$ are in a wave gauge on the same neighbourhood, hence on the same (small enough) strip, and satisfy the reduced Einstein equations. They coincide in a neighbourhood of M_0, since they take the same Cauchy data $g(.,0)$ and $\partial_t g(.,0)$, if they belong to $\tilde{E}_s$, $s > \frac{n}{2} + 1$. It will be the case only if the original metric is in $\tilde{E}_s$, $s > \frac{n}{2} + 2$, i.e. $s \geq 4$ if $n = 3$. □

Note that an $\tilde{E}_s(T)$ metric, $s > \frac{n}{2} + 2$ is in $C^2(T)$. Its geodesics are locally unique. Uniqueness has recently been proved to hold also for $s > 1 + \frac{n}{2}$ by I. Rodnianski (2008, personal communication) by using special properties of the wave equation on spacetimes with vanishing Ricci tensor.

8.4 *Causality*

The previous theorems are local in time, global in space. The next theorem, of great physical importance, can be interpreted as a proof of the propagation of gravitation with the speed of light.

We have the following definition (Appendix III, Definition 2.8).

Definition 8.7 *A (future) local* ***causal subset*** *of V based on $\Sigma_0 \subset M_0$ is a Stokes regular*[32] *compact subset of V, with non-empy interior and with boundary $\partial\Omega$ composed of three pieces $\Sigma_0 \subset M_0$, $\Sigma_\tau \subset M_\tau, 0 < \tau < T$ and L, the lateral boundary, an n-dimensional spacelike, or null and ingoing, submanifold.*

We now show that the solution of Einstein's equations in a local causal subset depends only on the data on the basis Σ_0 of this subset.

Theorem 8.8 *Let $(M, \bar{g}_1, K_1)$ and $(M, \bar{g}_2, K_2)$ be two initial data sets which coincide on an open subset Σ of M. Let Ω_1 and Ω_2 be two local causal subsets based on the same subset Σ, subsets respectively of a development (V_1, g_1) of $(M, \bar{g}_1, K_1)$ and (V_2, g_2) of $(M, \bar{g}_2, K_2)$. Then, there exist causal subsets based on Σ, $W_1 \subset \Omega_1$ and $W_2 \subset \Omega_2$, which are isometric.*

[31] Notations in Appendix III.

[32] That is, such that the Stokes formula is applicable in this subset, for instance if it lies on one side of a piecewise C^1 boundary.

Proof We have seen that the two solutions (V_1, g_1) and (V_2, g_2) admit representatives in a wave gauge respectively in neighbourhoods U_1 and U_2 of $M \times \{0\}$ in $M \times R$. We deduce also from Theorem 8.4 that if the initial data for g_1 and g_2 coincide on Σ it is possible to choose the wave gauges so that the initial values for the representatives coincide on Σ. The causality Theorem 2.12 in Appendix III, proved for second-order quasidiagonal hyperbolic systems, shows that the two solutions in wave gauge coincide on any local causal subset of the representative of g_1 based on Σ. □

Our definition of causal subsets, linked with the slicing of the manifold, is not invariant by isometries; however, the image under such isometries of a causal subset always includes a causal subset.

The theorem permits us to formulate the local existence and uniqueness theorem for Cauchy data in the local spaces $H_s^{u.l}$ defined in Appendix III. We formulate this theorem for the geometric Einstein equations.

Theorem 8.9 *(Local in time existence and uniqueness).*
Hypothesis on the initial data set $(M, \bar{g}, K)$.

1. (M, e) *is a smooth Sobolev regular Riemannian manifold.*
2. $\bar{g}$ *is a continuous properly Riemannian metric on* M *uniformly equivalent to* e *and such that:*

$$\bar{g} \in H_s^{u.l}, \quad K \in H_{s-1}^{u.l}.$$

Conclusions. If $s > \frac{n}{2} + 1$,
the initial data set admits an Einsteinian (vacuum) development (V_T, g), $V_T \equiv M \times [0, T)$, *such that the spacetime metric* g *is*[33] *a regularly sliced Lorentzian metric and the induced metric and extrinsic curvature of space slices belong respectively to* $E_s^{u.l.}(T)$ *and* $E_{s-1}^{u.l.}$, *where*

$$E_s^{u.l.}(T) := \bigcap_{0 \le k \le s} C^k([0, T], H_{s-k}^{u.l}).$$

Let (V_{T_1}, g_1) *and* (V_{T_2}, g_2) *be two such solutions. There are neighbourhoods of* M, $U_1 \subset V_{T_1}$ *and* $U_2 \subset V_{T_2}$, *such that two such solutions* (U_1, g_1) *and* (U_2, g_2) *are isometric in a causal subset based on* Σ *if* g_1 *and* g_2 *take the same initial values* $(\bar{g}, K)$ *on* Σ *and* $s > \frac{n}{2} + 2$, *even if* $s > \frac{n}{2} + 1$ *(Planchon and Rodnianski).*

We will return to causality and the global aspects of the Cauchy problem in Chapter 12.

[33] Recall that (see Appendix III) $\tilde{\mathrm{E}}_s(T)$ is the space of tensor fields u such that $u \in \bar{C}(V_T)$, space of continuous and bounded functions on V_T, while $\partial_t u$, $Du \in E_{s-1}(T) \equiv \bigcap_{k \le s-1} C^k([0, T), \overset{0}{H}_{s-k-1})$.

9 Constraints in a wave gauge

We now turn to consideration of the constraints satisfied by initial data, they are not a dynamical evolution system.

In wave coordinates the constraints read

$$S^{\alpha 0}_{(h)} + \frac{1}{2} g^{00} \partial_t F^\alpha - T^{\alpha 0} = 0. \tag{9.1}$$

They do not contain any second derivative of the metric with respect to time. The expressions of $S^{\alpha 0}$ and F^α show that these constraints are of the form (recall that $\mathcal{G}^{\alpha\beta}$ denote a tensor density; see (7.18) and (7.19)

$$\mathcal{C}^\alpha \equiv \frac{1}{2} g^{ij} \partial^2_{ij} \mathcal{G}^{\alpha 0} - \frac{1}{2} g^{00} \partial^2_{it} \mathcal{G}^{\alpha i} + g^{i0} \partial^2_{it} \mathcal{G}^{\alpha 0} + \mathcal{K}^{\alpha 0} - T^{\alpha 0} = 0, \tag{9.2}$$

where $\mathcal{K}^{\alpha 0}$ depends only on $\mathcal{G}$ and its first derivatives.

Remark 9.1 *Giving $\mathcal{G}^{\alpha\beta}, \partial_t \mathcal{G}^{\alpha\beta}$ for $t = t_0$ is equivalent to giving $g_{\alpha\beta}$, $\partial_t g_{\alpha\beta}$, for $t = t_0$.*

The constraints appear as a system of $n + 1$ equations for $(n + 1)(n + 2)$ unknowns, $\mathcal{G}^{\alpha\beta}, \partial_t \mathcal{G}^{\alpha\beta}$; hence they constitute an undetermined system. It is natural to try to split the initial data into specified quantities and $n+1$ unknowns, for which we wish to find a well-posed elliptic system. This was done for the first time in the generic case[34] by using Equations (9.2). Two choices have been suggested for splitting the initial data.

1. Give arbitrarily on the initial manifold M the quantities $\mathcal{G}^{\alpha i}$. The $\partial_t \mathcal{G}^{\alpha 0}$ are then determined by the harmonicity conditions (7.2). The unknowns are $\mathcal{G}^{00}$ and $\partial_t \mathcal{G}^{ij}$. Equations (9.2) can be written in the form, for $\alpha = 0$,

$$\mathcal{C}^0 \equiv \frac{1}{2} g^{ij} \partial^2_{ij} \mathcal{G}^{00} - A = 0, \tag{9.3}$$

where

$$A := \frac{1}{2} g^{00} \partial^2_{it} \mathcal{G}^{\alpha i} + g^{i0} \partial^2_{it} \mathcal{G}^{\alpha 0} + \mathcal{K}^{\alpha 0} - T^{\alpha 0} \tag{9.4}$$

depends only on the specified quantities and on the first derivatives of the unknowns, and for $\alpha = h$

$$\mathcal{C}^h \equiv -\frac{1}{2} g^{00} \partial_i (\partial_t \mathcal{G}^{hi}) + B^h = 0, \tag{9.5}$$

where

$$B^h := \frac{1}{2} g^{ij} \partial^2_{ij} \mathcal{G}^{h0} + g^{i0} \partial^2_{it} \mathcal{G}^{h0} + \mathcal{K}^{\alpha 0} - T^{\alpha 0}. \tag{9.6}$$

[34] (Choquet-)Bruhat, Y. (1962) "The Cauchy problem" in *Gravitation: An Introduction to Current Research* (ed. L. Witten), pp. 130–68, Wiley.

Equations (9.5) can be put into elliptic form using a procedure analogous to the one which is discussed for the momentum constraint in Chapter 7 for the case of the conformal method.

2. In the spirit of Wheeler's "thin sandwich conjecture", give arbitrarily on M the quantities $\mathcal{G}^{ij}$. The harmonicity conditions then determine $\partial_t \mathcal{G}^{0j}$. Give also arbitrarily $\partial_t \mathcal{G}^{ij}$. Use again the harmonicity condition to write

$$\partial_t \mathcal{G}^{00} = -\partial_i \mathcal{G}^{0i}. \tag{9.7}$$

The constraints then become an elliptic system for the unknowns $\mathcal{G}^{00}$, $\mathcal{G}^{0i}$, which reads

$$\mathcal{C}^0 \equiv \frac{1}{2} g^{ij} \partial^2_{ij} \mathcal{G}^{00} - g^{0i} \partial^2_{ij} \mathcal{G}^{j0} - f^0 = 0, \tag{9.8}$$

$$\mathcal{C}^h \equiv \frac{1}{2} g^{ij} \partial^2_{ij} \mathcal{G}^{0h} - f^h = 0, \tag{9.9}$$

where f^0 and f^h depend only on the given quantities and on the first derivatives of the unknowns.

The linearization of Equations (9.8) and (9.9) in Euclidean space reduces to the Euclidean Laplace equations. They have been shown by Vaillant-Simon[35] to have asymptotically Euclidean solutions in a neighbourhood of Euclidean space.

One can easily extend the above construction for the constraints to the global in space case of $\hat{e}$ wave gauge: the partial derivatives ∂_α have then to be replaced by $\hat{e}$-covariant derivatives D_α.

We will return to the analysis of the constraints in the next chapter, in a gauge-independent context. We remark however that the equations given here may be relevant for numerical study, in particular for evolution computed in a wave gauge. Indeed it has been found recently, to no one's surprise, that the computation of solutions in a wave gauge has a better stability behaviour than the ADM flow. In fact, the better behaviour is obtained by using a damped wave gauge (see Section 7.3), which apparently damps the error.

10 Einstein equations with field sources

Scalar field or wave map sources satisfy linear or semilinear wave equations in the spacetime metric g. They are hyperbolic and causal provided that this metric is Lorentzian. The coupling of these equations to the reduced Einstein equations in a wave gauge results in a system of quasilinear wave equations, obviously hyperbolic for Lorentzian g. The local existence and uniqueness theorems proved for the vacuum Einstein equations can be easily extended to Einstein equations with scalar field or wave map sources. We leave to the reader the wording of these theorems.

[35] Vaillant-Simon, A. (1969) *J. Math. Pures Appl.*, **48**.

On the other hand, the classical Yang–Mills sources (in particular, electromagnetic ones) split like the Einstein equations into constraints and evolution equations. These evolution equations are hyperbolic in a well-chosen gauge. We give the proof below for the electromagnetic field.

Recall that the Einstein equations with electromagnetic source read (see Chapter 3)

$$S_{\alpha\beta} = \tau_{\alpha\beta} \equiv F_{\alpha}{}^{\lambda} F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}, \tag{10.1}$$

or equivalently

$$R_{\alpha\beta} = F_{\alpha}{}^{\lambda} F_{\beta\lambda} - \frac{1}{2(n-1)} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}. \tag{10.2}$$

The first set of Maxwell's equations is

$$dF \equiv 0 \quad \text{i.e.} \quad \nabla_{\alpha} F_{\beta\gamma} + \nabla_{\gamma} F_{\alpha\beta} + \nabla_{\beta} F_{\gamma\alpha} = 0. \tag{10.3}$$

The second set of Maxwell's equations *in vacuo* is

$$\delta F = 0, \quad \text{i.e. } \nabla_{\alpha} F^{\alpha\beta} = 0. \tag{10.4}$$

10.1 *Maxwell constraints*

Given an n-dimensional submanifold $M_0 := M \times \{t_0\}$ of a spacetime $(V = M \times R, g)$. The electromagnetic initial data on M_0 are a 2-form $\bar{F}$ and a vector field $\bar{E}$. The 2-form $\bar{F}$ is the form induced on M_0 by the electromagnetic field F, while $\bar{E}$ is the electric vector field on M_0 relative to the unit normal to M_0 in the spacetime metric. In a Cauchy-adapted frame the components of $\bar{E}$ are, denoting by an overbar values taken on M_0,

$$\bar{E}^i := \bar{N}^{-1} \bar{F}^i_0. \tag{10.5}$$

The Maxwell constraints are the part of the Maxwell equations which do not depend on time derivatives of $\bar{F}$ or $\bar{E}$. They are

$$d\bar{F} = 0 \quad \text{i.e} \quad \bar{\nabla}_h \bar{F}_{ij} + \bar{\nabla}_j \bar{F}_{hi} + \bar{\nabla}_i \bar{F}_{jh} = 0 \tag{10.6}$$

and

$$\overline{\nabla_{\alpha} F^{\alpha 0}} = 0, \tag{10.7}$$

which can be shown to be equivalent to

$$\bar{\nabla}_i \bar{E}^i = 0. \tag{10.8}$$

We will in the next section use the locally defined vector potential A, introduced in Chapter 3 such that $F = dA$. Then $\bar{F} = d\bar{A}$, the first Maxwell constraint (10.6) is automatically satisfied.

10.2 *Lorentz gauge*

In a simply connected domain general solution of the first set of Maxwell equations $dF = 0$, is

$$F = dA, \quad \text{i.e.} \quad F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha, \tag{10.9}$$

where A is the vector potential 1-form. The second set (*in vacuo*) is

$$\delta F = 0, \quad \text{i.e.} \quad \nabla^\alpha F_{\alpha\beta} = g^{\alpha\lambda}(\partial_\lambda F_{\alpha\beta} - \Gamma^\mu_{\lambda\beta} F_{\alpha\mu}) - g^{\alpha\lambda}\Gamma^\mu_{\lambda\alpha} F_{\mu\beta} = 0;$$

It reduces in wave coordinates; that is when[36]

$$\Gamma^\mu := g^{\alpha\lambda}\Gamma^\mu_{\lambda\alpha} = 0, \tag{10.10}$$

to a system which reads in terms of A

$$g^{\alpha\lambda}\{\partial_\lambda\partial_\alpha A_\beta - \partial_\lambda\partial_\beta A_\alpha - \Gamma^\mu_{\lambda\beta}(\partial_\alpha A_\mu - \partial_\mu A_\alpha) = 0.$$

By elementary manipulations these equations become

$$g^{\alpha\lambda}\partial_\lambda\partial_\alpha A_\beta - \partial_\beta(g^{\alpha\lambda}\partial_\lambda A_\alpha) + (\partial_\beta g^{\alpha\lambda})\partial_\lambda A_\alpha - \Gamma^\mu_{\lambda\beta}(\partial_\alpha A_\mu - \partial_\mu A_\alpha) = 0. \tag{10.11}$$

They reduce to a quasidiagonal semilinear system of wave equations for the components of A, of the form

$$g^{\alpha\lambda}\partial_\lambda\partial_\alpha A_\beta = f_\beta(g, \partial g, \partial A), \tag{10.12}$$

if A satisfies the gauge condition

$$g^{\alpha\lambda}\partial_\lambda A_\alpha \equiv g^{\alpha\lambda}(\nabla_\lambda A_\alpha + \Gamma^\mu_{\alpha\lambda} A_\mu) = 0, \tag{10.13}$$

which is equivalent in wave coordinates to the Lorentz gauge condition

$$\delta A = 0, \quad \text{i.e.} \quad \nabla_\alpha A^\alpha = 0. \tag{10.14}$$

Equations (10.13) together with the Einstein equations in wave gauge, with source F, constitute a quasidiagonal quasilinear system of wave equations for the pair g, A. Local existence and uniqueness for a solution of the Cauchy problem for the pair (g, A) in wave and Lorentz gauge results from the general theorems of Appendix III, Section 4.3.

Various other gauge conditions can be used on A to solve Maxwell's equations: temporal gauge $A_0 = 0$, Coulomb gauge $\nabla_i A^i = 0$, and Cromström gauge[37]. They have proved useful in different domains.

All these gauges, like the Lorentz gauge, generalize to Yang–Mills fields. They lead to equations with the same principal parts as in the case of electromagnetism.

[36] We denote Γ the gauge functions previously denoted F to avoid confusion with the electromagnetic field.

[37] See Eardley, D. and Moncrief, V. (1982) *Commun. Math. Phys.*, **83**, 171–212, and Chruściel, P. and Shatah, J. (on a Lorentzian manifold).

10.3 *Existence and uniqueness theorems*

To show that the pair (g, A) obtained in wave and Lorentz gauge satisfies the full set of Einstein–Maxwell equations, when the initial data satisfy the Einstein–Maxwell constraints, we follow the same steps as in the vacuum case.

Lemma 10.1 *If the pair (g, A) satisfies the Einstein–Maxwell equations in the wave and Lorentz gauges, the wave functions Γ^α and the function δA satisfy a quasidiagonal system of linear homogeneous differential equations with principal terms the wave equation in the metric g.*

Proof It uses the Bianchi identities and the identities $d^2 \equiv 0$ and $\delta^2 \equiv 0$ satisfied by the differential and codifferential operators d and δ. □

Exercise. *Write up the proof.*

Lemma 10.2 *For a solution of the wave and Lorentz gauge reduced Einstein–Maxwell equations, with initial data satisfying $\Gamma^\alpha|_M = 0$, $\delta A|_M = 0$, the conditions $\partial_t \Gamma^\alpha|_M = 0$, $\partial_t \delta A|_M = 0$ are satisfied if and only if the initial data satisfy the Einstein–Maxwell constraints.*

Proof Straightforward calculation. □

It results from these two lemmas that a solution for the reduced Einstein–Maxwell equations in wave and Lorentz gauge with initial data satisfying the Einstein–Maxwell constraints, is a solution for the full Einstein–Maxwell system, if the functional spaces corresponding to this solution are such that the uniqueness theorem of Appendix III applies to the linear equations satisfied by Γ and δA.

Since the results and their proofs for the Einstein–Maxwell equations are similar to those for the vacuum Einstein system, we state directly the strongest result, analogous to Theorem 8.9.

Theorem 10.3 *Let $(M; \bar{g}, K, \bar{F}, \bar{E})$ be an Einstein–Maxwell initial data set satisfying the Einstein and Maxwell constraints. Suppose that $(M, \bar{g}, K)$ satisfy the hypothesis of Theorem 8.9 for vacuum, and that $\bar{F} \in H^{u.l}_{s-1}$, $\bar{E} \in H^{u.l}_{s-1}$, $s > \frac{n}{2} + 1$. This initial data set admits an Einstein–Maxwell development, $(V_T; g, F)$ such that $V_T \equiv M \times (-T, T)$, the spacetime metric g is a regularly sliced Lorentzian metric, the induced metric on space slices belongs to $E^{u.l.}_s(T)$. The extrinsic curvature of space slices as well as the 2-form induced by F and the electric field associated to F on space slices belong $E^{u.l.}_{s-1}$.*

Proof To the geometric initial data one associates standard PDF initial data for g as in the vacuum case (Section 7.5), using the harmonicity conditions $\bar{\Gamma} = 0$. To $\bar{F}$, satisfying the Maxwell constraint $d\bar{F}$, we associate an initial value $\bar{A}$ of the vector potential A, such that $d\bar{A} = \bar{F}$, locally determined up to the differential of a function ϕ. The value $\bar{A}_0$ of the time component of A on M, $\bar{A}_0$ is another

arbitrary quantity. The vector $\bar{E}_i$ determines then $\overline{\partial_t A_i}$, and the Lorentz gauge condition (11.14) taken on M determines $\overline{\partial_t A_0}$. □

Corollary 10.4 *The solution (g, F) is locally geometrically unique if $s > \frac{n}{2}+2$.*

Proof We leave to the reader the proof that it does not depend on the choice of $\bar{A}$ and $\overline{A_0}$, given $\bar{F}$. □

In the case of Yang–Mills field the potential is part of the original equations. The Cauchy problem is still locally well posed[38], but the formulation and proof are a bit more complicated.

10.4 *Wave equation for F*

10.4.1 *Case of electromagnetism*

In the case of electromagnetism, the equations do not require the introduction of an electromagnetic potential A: the Maxwell equations involve only the 2-form F which satisfies the equations $dF = 0$ and $\delta F = 0$. It results from these equations that F satisfies a quasidiagonal, quasilinear system

$$(\delta d + d\delta)F = 0; \tag{10.15}$$

that is[39]

$$g^{\alpha\beta}\nabla_\alpha\nabla_\beta F_{\lambda\mu} + R^\alpha_\mu F_{\alpha\lambda} - R^\alpha_\lambda F_{\alpha\mu} + 2R^\alpha{}_{\lambda,}{}^\beta{}_\mu F_{\alpha\beta} = 0. \tag{10.16}$$

To apply general theorem for quasilinear, quasidiagonal second-order systems to this system coupled with the Einstein equations in wave gauge one uses the Leray–Volevic indices (see Appendix IV) to define the principal parts, because the above equation contains second derivatives of g.

Lemma 10.5 *The wave equation for F together with the Einstein equations in a wave gauge with source the Maxwell stress energy tensor are a Leray hyperbolic system of quasidiagonal, quasilinear wave equations for g and F.*

Proof We give to the unknowns and to the equations the Leray–Volevic weights

$$m(g) = 3, \quad m(F) = 2 \quad \text{and} \quad n(Einstein) = 1, \quad n(wave \text{ for } F) = 0 \tag{10.17}$$

The principal matrix is then diagonal with diagonal term the wave operator. The Leray theorem gives the result for the reduced Einstein and equation for F system. □

The wave equation for F and the choice of a wave gauge for the metric (defined through a wave map), allows for the global formulation in space without restriction on the topology of the initial manifold M.

[38] Choquet-Bruhat, Y. (1991) Marcel Grossmann meeting.

[39] See for instance the expression of this operator for an arbitrary p-form in CB-DM1 V B 4.

Theorem 10.6 *Let $(M, \bar{g}, K, \bar{F}, \bar{E})$ be an initial data set satisfying the Einstein and Maxwell constraints. Suppose that $(M, \bar{g}, K)$ satisfy the hypothesis of Theorem (7.4) for vacuum, and that $\bar{F} \in H_{s-1}$, $\bar{E} \in H_{s-1}$, $s > \frac{n}{2} + 1$. Then, the Einstein–Maxwell equations have a solution (V_T, g, F), $g \in \tilde{E}_s$, $F \in E_{s-1}$, with (V, g) a causal Einsteinian development of $(M, \bar{g}, K)$ with source the Maxwell tensor. This development is unique in the class of maximal globally hyperbolic developments if $s > \frac{n}{2} + 2$.*

10.4.2 *Case of Yang–Mills fields*

A Yang–Mills field F on spacetime also satisfies a wave type equation, but only semilinear. Indeed the Yang–Mills gauge covariant derivative is defined by

$$\hat{\nabla} F := \nabla \bar{F} + [A, F], \tag{10.18}$$

with ∇ the Riemannian covariant derivative and $[.,.]$ the bracket in the Lie algebra corresponding to the considered Yang–Mills model. The Yang–Mills equations *in vacuo* are

$$\hat{\nabla}_\alpha F_{\beta\gamma} + \hat{\nabla}_\gamma F_{\alpha\beta} + \hat{\nabla}_\beta F_{\gamma\alpha} = 0 \tag{10.19}$$

and

$$\hat{\nabla}_\alpha F^{\alpha\beta} = 0. \tag{10.20}$$

These equations imply the second-order semilinear equation for F, depending on A,

$$\hat{\nabla}^\lambda \hat{\nabla}_\lambda F_{\alpha\beta} - 2[F^\gamma{}_{,\alpha}\, , F_{\gamma\beta}] = 0 \tag{10.21}$$

where

$$\hat{\nabla}^\lambda \hat{\nabla}_\lambda F_{\alpha\beta} \equiv \nabla^\lambda \nabla_\lambda F_{\alpha\beta} + 2\nabla_\gamma [A^\gamma, F_{\alpha\beta}] - [\nabla_\gamma A^\gamma, F_{\alpha\beta}] + [A_\gamma, [A^\gamma, F_{\alpha\beta}]. \tag{10.22}$$

Exercise. *Prove this equation, using the curvature tensor of the gauge connection.*

This equation for F plays an essential role in the proof of the global existence of a solution of the Cauchy problem for the Yang–Mills equations by Eardley and Moncrief[40].

[40] Eardley, D. and Moncrief, V. (1982) *Commun. Math. Phys.*, **83**, 171–212. See also the survey article Choquet-Bruhat, Y. (1983) in *Relativity, Cosmology, Topological Mass and Supergravity* (ed. C. Aragone), pp. 108–35, World Scientific.

VII

CONSTRAINTS

1 Introduction

We have seen (Section VI.4) that the geometric initial data on an n-dimensional manifold M for the Einstein equations are $\bar{g}$ a properly Riemannian metric and K a symmetric 2-tensor[1]. They cannot be arbitrary, since they must satisfy the constraints. These are equations on M given by:

The **Hamiltonian constraint**

$$\mathcal{H}(\bar{g}, K) \equiv R(\bar{g}) - |K|^2_{\bar{g}} + (tr_{\bar{g}} K)^2 - 2\rho = 0, \tag{1.1}$$

The **momentum constraint**:

$$\mathcal{M}(\bar{g}, K) \equiv \bar{\nabla}.K - \bar{\nabla} tr K - J = 0, \tag{1.2}$$

ρ is a scalar and J a vector on M determined by the projection on M and the normal to M, when embedded in a spacetime (V, g), of the stress energy tensor T of the sources. In a Cauchy-adapted frame, where the equation of M in V is $x^0 \equiv t = 0$ one has

$$J^i = NT^{0i}, \quad \rho = N^2 T^{00}, \tag{1.3}$$

where N is the lapse associated with the embedding. It holds that $\rho \geq 0$ for usual sources.

When the sources ρ and J are known, the unknowns in the constraints (1.1) and (1.2) are the metric $\bar{g}$ and the tensor K. It is mathematically clear that these scalar and vector equations do not have a unique solution, even geometrically. Physically this fact corresponds to the property of Einsteinian gravity being its own source. Roughly speaking "radiation data" should also be given. Modulo such additional arbitrary data the Einstein constraints should determine the full Cauchy data on an initial spacelike manifold. This strongly suggests the search for the replacement of the constraints, modulo well-chosen arbitrary data, by an elliptic system to be satisfied by a scalar and a vector fields.

We have in Section VI.7 written the constraints as such an elliptic second-order system in a thin sandwich formulation by using wave coordinates and taking as arbitrarily given the tensor densities $\mathcal{G}^{ij}$ and $\partial_t \mathcal{G}^{ij}$. We have obtained solutions

[1] When we work only on one spacelike n-manifold M, we usually suppress the overbar on quantities defined on M, K, ρ, $J, \ldots$, except for $\bar{\mathrm{g}}$, the induced metric, and geometric quantities defined through it.

in a neighbourhood of Minkowskian data. In this chapter we will take greater advantage of the geometric character of the constraints.

In Section 2 we study the set of solutions, i.e. linearization of the constraints and stability of solutions of the full constraints, for given sources; we show that this set of solutions is a smooth infinite dimensional manifold, except at some exceptional points.

In the subsequent sections we cast the constraints into the form of a second-order elliptic system, here using the conformal method. The geometric meaning of the lower-order terms permits the discussion of existence and uniqueness of solutions on general manifolds, without restriction to the neighbourhood of a known solution.

We will treat two cases, using the definitions and theorems of Appendix II.

1. The case where M is a compact manifold, which we endow with a smooth Riemannian metric e in order to define Sobolev spaces W^p_s. We denote by M^p_s the space of Riemannian metrics $\bar{g} \in W^p_s$. This is an open cone in W^p_s if $s > \frac{n}{p}$.

2. The case where $(M, \bar{g})$ is asymptotically Euclidean. We use then the weighted Sobolev spaces $W^p_{s,\delta}$. The space $M^p_{s,\delta}$ is the space of Riemannian metrics such that $\bar{g} - e \in W^p_{s,d}$, where e is a smooth metric on M euclidean at infinity. If $s > \frac{n}{p}$, $\delta > -\frac{n}{p}$, it is an open set in $W^p_{s,\delta}$ of metrics tending pointwise to the flat metric at infinity.

The case of asymptotically hyperbolic data has been treated by Andersson, Chruściel, and Friedrich[2] and by Isenberg and Park[3]. Other, very general, cases have been treated by Chruściel and Delay[4].

Unless otherwise stated we suppose *the initial manifold to be oriented*.

2 Linearization and stability

The concept of linearization stability was introduced by Choquet-Bruhat and Deser[5] with application to Minkowskian initial data. It was extended to compact manifolds by O'Murchad and York, and with another method by Fisher and Marsden. Their method was extended to asymptotically Euclidean manifolds, in the case $TrK = 0$, by Choquet-Bruhat, Fisher, and Marsden and in the general case, with a partly heuristic proof, by Beig and O'Murchada. A rigorous proof of the surjectivity of the linearization of the constraints in the asymptotically Euclidean case when $TrK \neq 0$ was finally given by Corvino and Schoen[6].

[2] Andersson, L., Chruściel, P., and Friedrich, H. (1992) *Commun. Math. Phys.*, **149**, 587–612.

[3] Isenberg, J. and Park (1996) gr.qc 9610027 Vl.

[4] Chruściel, P. and Delay, E. (2003) *Mémoire Soc. Math. France*, **94**.

[5] Choquet-Bruhat, Y. and Deser, S. (1972) *C. R. Acad. Sci. Paris*, **274**, 682–4. See also *Ann. Phys.*, **81**, 165–78, 1973.

[6] Corvino, M. and Schoen, R. (2006) *J. Diff. Geom.*, **73**(3), 185–217.

For simplicity we consider essentially the vacuum case. The extension to the non-vacuum case presents no conceptual difficulty.

We have seen that a development of an initial data set $(M, \bar{g}, K)$ depends continuously on $\bar{g}$ and K, in some Banach space B of symmetric 2-tensors on M adapted to the considered problem. The problem is now: given some initial data $(\bar{g}_0, K_0) \in B$, do there exist nearby initial data $(\bar{g}, K)$ on M which also satisfy the constraints? It is clear that such data cannot fill an open set of the whole space B, because the tangent to a C^1 curve of initial data passing through $(\bar{g}_0, K_0)$ is not arbitrary: it must satisfy the linearized constraints. Such a tangent is given by a pair of symmetric 2-tensors $(h, k) \in B$ such that

$$\Phi_0'(h, k) = 0, \tag{2.1}$$

where Φ_0' denotes the functional (Fréchet) derivative at $(\bar{g}_0, K_0)$ of the constraint map

$$\Phi : \quad (\bar{g}, K) \mapsto (\mathcal{H}(\bar{g}, K), 2\mathcal{M}(\bar{g}, K)). \tag{2.2}$$

The problem of stability of a solution $(\bar{g}_0, K_0)$ of the constraints is therefore: to each solution (h, k) of the linearized constraints (2.1) does there correspond a C^1 curve of solutions of the non-linear constraints with tangent (h, k) at $(\bar{g}_0, K_0)$? Equivalently, is there a subset $\{\Phi(\bar{g}, K) = 0\} \cap U$ which is a submanifold of B modelled on the vector space defined by the linearized constraint (2.1). The solution $(\bar{g}_0, K_0)$ is said to be **linearization stable**[7] if the answer is yes. It is known from a counterexample of Brill and Deser, the flat 3-torus, that solutions are not always linearization stable.

2.1 *Linearization of the constraints map, adjoint map*

In this subsection, to make the reading easier, we give formal computations, which are mathematically justified in the situations that we will examine later.

The derivative Φ' of the constraints map Φ is somewhat easier to write if considered as a map from $\bar{g}$ and the mixed tensor form of K, into the product of a space of functions by a space of vector fields. Then

$$\mathcal{H}(\bar{g}, K) \equiv \bar{R} - |K|^2_{\bar{g}} + (tr_{\bar{g}} K)^2, \tag{2.3}$$

with

$$\bar{R} := R(\bar{g}), \qquad |K|^2_{\bar{g}} := K^i_j K^j_i, \qquad tr_{\bar{g}} K := K^i_i.$$

$$\mathcal{M}(\bar{g}, K) \equiv \bar{\nabla} K - \partial tr K, \quad \text{i.e.} \quad \mathcal{M}_i(\bar{g}, K) \equiv \bar{\nabla}_j K^j_i - \partial_i (tr K). \tag{2.4}$$

The derivative Φ' at a pair $(\bar{g}, K)$ is the linear mapping Φ' acting on a pair (h, k) given by:

$$\Phi' : (h, k) \mapsto \delta\Phi \equiv (\delta\mathcal{H}, 2\delta\mathcal{M}), \tag{2.5}$$

[7] Terminology due to A. Fisher and J. Marsden, who introduced the surjectivity criterion.

with[8]

$$\delta\mathcal{H} = \delta\bar{R} - 2k.K + 2(tr_{\bar{g}}k)(tr_{\bar{g}}K) \tag{2.6}$$

where by straightforward computation, we find that

$$\delta\bar{R} = -\Delta_{\bar{g}}tr_{\bar{g}}h + div_{\bar{g}}(div_{\bar{g}}h) - h.Ricci(\bar{g}), \tag{2.7}$$

with

$$div_{\bar{g}}(div_{\bar{g}}h) \equiv \bar{\nabla}^i\bar{\nabla}^j h_{ij}$$

and

$$\delta\mathcal{M}_i = \bar{\nabla}_j k_i^j - \partial_i(trk) + \frac{1}{2}K_i^l\partial_l tr_{\bar{g}}h - \frac{1}{2}K^{jl}\bar{\nabla}_i h_{jl}. \tag{2.8}$$

We have seen in Section VI.5.2 that the formal L^2 adjoint Φ'^* of the operator Φ' is a mapping from a pair (f, X), with f a function and X a vector field, into a pair of 2-tensor fields denoted $\Phi'^*_{\bar{g}}(f, X)$ and $\Phi'^*_P(f, X)$:

$$\Phi'^* : (f, X) \mapsto \Phi'^*(f, X) \equiv (\Phi'^*_{\bar{g}}(f, X), \Phi'^*_P(f, X)) \tag{2.9}$$

These tensors are such that

$$\langle (h, k), \Phi'^*(f, X)\rangle = \langle \Phi'(h, k), (f, X)\rangle; \tag{2.10}$$

that is (a dot is a scalar product in the metric $\bar{g}$),

$$\int \{h.\Phi'^*_{\bar{g}}(f, X) + k.\Phi'^*_K(f, X)\}d\mu_{\bar{g}} = \int \{f\delta\mathcal{H} + 2X^i\delta\mathcal{M}_i\}\mu_{\bar{g}} \tag{2.11}$$

The expressions of $\delta\mathcal{H}$ and $\delta\mathcal{M}$, and integration by parts, show that the values of $\Phi'^*_{\bar{g}}(f, X)$ and $\Phi'^*_K(f, X)$ are given by

$$\Phi'^*_{\bar{g}}(f, X)^{ij} = -g^{ij}\Delta_{\bar{g}}f + \bar{\nabla}^i\bar{\nabla}^j f - \bar{R}^{ij}f - g^{ij}\bar{\nabla}_h(X^k K_k^h) + \bar{\nabla}_h(X^h K^{ij}) \tag{2.12}$$

and

$$\Phi'^*_K(f, X)^i_j \equiv -2f(K^i_j - \delta^i_j K^h_h) - 2\bar{\nabla}_j X^i + 2\delta^i_j\bar{\nabla}_h X^h. \tag{2.13}$$

Lemma 2.1 *The operator $\Phi'\Phi'^*$ is elliptic*[9].

Proof The principal part of the operator Φ'^* is obtained by giving Leray–Volevic–Nirenberg weights (integral numbers; see Section 2 in Appendix IV) to the equations and variables; we choose $m(f) = 2$, $n(X) = 2$ and $n(2.12) = 0$, $n(2.13) = 1$. The principal part acting on a pair (f, X) is a rectangular matrix with elements terms of order $m(f) - n(2.12)$, and so on. It is therefore

$$\begin{pmatrix} -g^{ij}\Delta_{\bar{g}}f + \bar{\nabla}^i\bar{\nabla}^j f & 0 \\ 0 & -2\bar{\nabla}_j X^i + 2\delta^i_j\bar{\nabla}_h X^h \end{pmatrix} \tag{2.14}$$

[8] In physicists' language $\delta\Phi$ is the linearization of Φ at ($\bar{g}$,K).
[9] Remark due to Chruściel and Delay (2003).

The principal symbol A^* of Φ'^*, acting on the pair (f, X) is the rectangular matrix

$$A^* \equiv \begin{pmatrix} (-g^{ij}\xi^h\xi_h + \xi^i\xi^j)f & 0 \\ 0 & -2\xi_j X^i + 2\delta^i_j \xi_h X^h \end{pmatrix}. \tag{2.15}$$

This matrix is injective for all $\xi \not\equiv 0$, if $n > 1$, since $(-g^{ij}\xi^h\xi_h + \xi^i\xi^j)f = 0$ implies by contraction $(-n+1)\xi^i\xi_i f = 0$, while $-2\xi_j X^i + 2\delta^i_j\xi_h X^h = 0$ implies by contraction $\xi_h X^h = 0$; hence $\xi_j X^i = 0$, and finally $X = 0$ if $\xi \neq 0$.

The principal symbol of the operator $\Phi'\Phi'^*$ is the product AA^*, with A, principal symbols of Φ', the adjoint matrix of A^*. Elementary formulae show that AA^* is a square matrix, injective like A^*, hence invertible; by definition, $\Phi'\Phi'^*$ is an elliptic operator (see Appendix II). □

Lemma 2.2 *The kernel of Φ'^* is the space of spacetime vectors $X \equiv (X^i, X^0 = f)$ which are the value on the initial manifold M of Killing vector fields of the spacetime which has $(\bar{g}, K)$ as initial data*[10].

Proof We have deduced in Chapter 6 from the Hamiltonian form of the evolution equations that the time derivatives $\partial\bar{g}/\partial t$ and $\partial P^{ij}/\partial t$ are given by $\Phi'^*(N, \beta)$. □

2.2 *Linearization stability*

The proofs of linearization stability rely on the implicit function theorem. The statements of the theorems on linearization stability or non-stability, depend on the Banach space to which the solutions belong. We treat the case where M is compact (without boundary) or asymptotically Euclidean. Other cases have been considered. A very general study is given in Chruściel and Delay (2003).

2.2.1 *Case of compact M*

In the case where M is compact, the elliptic operator $\Phi'\Phi'^*$ obeys the Fredholm alternative, and the following lemma holds (see Appendix II):

Lemma 2.3 *If M is a compact manifold, then Φ is a mapping from $(\bar{g}, K) \in M^p_{s+2} \times W^p_{s+1}$ into W^p_s, if $p > \frac{n}{2}$, $s \geq 0$, and it holds that:*

$$W^p_s = range\Phi' + ker\Phi'^*. \tag{2.16}$$

also

$$W^p_s = E_1 + E_2, \quad with \quad E_1 := ker\Phi', \quad E_2 := range\Phi'^*. \tag{2.17}$$

From this lemma we will deduce the following theorem.

[10] Such vector fields are sometimes called KIDS (Killing Initial Data). They have been proved to be non-generic by Beig, R., Chruściel, P., and Schoen, R. (2005) *Ann. H. Poincaré*, **6**, 428–46.

Theorem 2.4 *(Moncrief) The constraint set* $(M, \bar{g}, K)$, $(\bar{g}, K) \in M_2^p \times W_1^p$, $p > \frac{n}{2}$, *is linearization stable if its Einsteinian vacuum development admits no Killing vector field.*

Proof The spaces E_1 and E_2 defined in 2.17 are complementary Banach subspaces of W_2^p. The mapping Φ is a mapping $(y_1, y_2) \in (\{E_1 \cap M_2^p\}, E_2)$ into L^p. The mapping Φ' is surjective on L^p by Lemmas 2.2 and 2.3. Its partial derivative Φ_2' has the same range as Φ' since the action of Φ' on E_1 is zero; hence Φ_2' is also surjective. It is injective, because the decomposition (2.16) implies $E_2 \cap ker\Phi' = 0$. The equation $\Phi(y_1, y_2) = 0$, $y = (\bar{g}, K)$, has by the implicit function theorem, a solution in a neighbourhood of a given solution $(\bar{g}_0, K_0)$. □

Necessary conditions for linearization stability in the case of initial data leading to spacetimes with a Killing vector field are obtained by using the second derivative of Φ. It leads to another **Moncrief theorem**, which we give after the statement of the following lemma.

Lemma 2.5 *(Taub) Let* (V, g) *be a sliced*[11] *spacetime satisfying the Einstein equations*

$$Einstein(g) := S(g) = 0 \tag{2.18}$$

and let $\chi := \delta g$ *be a solution of the linearized equations*

$$S_g'.\chi = 0. \tag{2.19}$$

Then if X *is a Killing vector field of* g, *the integral, with* $S_{g,g}''$ *the second derivative of* S *with respect to* g *and* n *the spacetime unit normal to* M_t.

$$\int_{M_t} X.S_{g,g}''.(\chi, \chi).n d\mu_{\bar{g}} \tag{2.20}$$

is invariant under the flow of X.

Proof Use the linearized Bianchi identities and their derivation, together with integration by parts. □

The integrals (2.20) are called **Taub conserved quantities**.

Theorem 2.6 *(Moncrief) A necessary condition for* $(\bar{g}, K)$ *to be linearization stable is that the Taub conserved quantities vanish for every Killing vector field of its vacuum Einsteinian development.*

Proof It uses[12], together with the expression of the second derivative Φ'', the fact that if $\lambda \mapsto u(\lambda) := (\bar{g}(\lambda), K(\lambda))$, is a curve of solutions of the constraints,

[11] Definition in Chapter 6.

[12] See details in Choquet-Bruhat, Y. and York, J. (1979) *General Relativity* (ed. A. Held), Plenum, New York, or in the original articles quoted there.

i.e. $\Phi(u(\lambda) = 0$, then the following equation holds

$$\frac{d^2}{d\lambda^2}\Phi(u(\lambda)) \equiv \Phi'(u)\frac{d^2u}{(d\lambda)^2} + \Phi''(u)\left(\frac{du}{d\lambda}, \frac{du}{d\lambda}\right) = 0. \tag{2.21}$$

□

2.2.2 *Case of M asymptotically Euclidean*

The **Fisher–Marsden criterion** for linearization stability is the surjectivity of the linear map Φ' whose kernel E_1 has a closed complement E_2. It is then possible to apply the implicit function theorem to the mapping Φ from $E_1 \times E_2$ onto F by proving that Φ'_2 is an isomorphism $E_2 \to F$.

This criterion can be used in the asymptotically Euclidean case to prove the following theorem.

Theorem 2.7 *The constraint set* $(M, \bar{g}, K)$, $(\bar{g}, K) \in M^p_{2,\delta} \times W^p_{1,\delta}$, $p > \frac{n}{2}$, *is linearization stable if* $\delta > -\frac{n}{p}$.

The proof is very simple when $trK = 0$. Indeed, the mapping $\Phi' : (h, k) \mapsto (\delta\mathcal{H}, \delta\mathcal{M})$ is surjective from $W^p_{2,\delta} \times W^p_{1,\delta+1}$ onto $(W^p_{0,\delta+2}, W^p_{0,\delta+2})$ if the equations

$$\delta\mathcal{H} \equiv -\Delta_{\bar{g}} tr_{\bar{g}} h + \bar{\nabla}^i\bar{\nabla}^j h_{ij} - \bar{R}^{ij}h_{ij} - 2k^j_i K^i_j + 2(trk)(trK) = f_0 \tag{2.22}$$

and

$$\delta\mathcal{M}_i \equiv \bar{\nabla}_j k^j_i - \partial_i(trk) + \frac{1}{2}K^l_i\partial_l tr_{\bar{g}} h - \frac{1}{2}K^{jl}\bar{\nabla}_i h_{jl} = f_i \tag{2.23}$$

have a solution $(h, k) \in W^p_{2,\delta} \times W^p_{1,\delta+1}$ for each pair $(f_0, f_i) \in (W^p_{0,\delta+2}, W^p_{0,\delta+2})$. Suppose $trK = 0$ and look for a solution of the form

$$h_{ij} = \frac{1}{n}\lambda\bar{g}_{ij}, \quad k^j_i = \frac{1}{2}\{-\lambda K^j_i + (\mathcal{L}_{conf,\bar{g}}X)^j_i\} \tag{2.24}$$

where $\mathcal{L}_{conf,\bar{g}}X$ is the conformal Lie derivative of $\bar{g}$ with respect to the vector field X. If $trK = 0$, then also $trk = 0$. Equation (2.22) reads for such a pair (h, k):

$$\delta\mathcal{H} = \left(\frac{1}{n} - 1\right)\Delta_{\bar{g}}\lambda - \lambda\left(\frac{1}{n}\bar{R} - |K|^2\right) = f_0 \tag{2.25}$$

i.e. if $(\bar{g}, K)$ satisfies the Hamiltonian constraint (with $trK = 0$):

$$\left(\frac{1}{n} - 1\right)(\Delta_{\bar{g}}\lambda - \lambda|K|^2) - K^i_j(\mathcal{L}_{conf,\bar{g}}X)^j_i = f_0 \tag{2.26}$$

and, after simplification using the momentum constraint ($\bar{\nabla}_i K^i_j = 0$, with $trK = 0$):

$$\delta\mathcal{M}_i = (\bar{\Delta}_{\bar{g},conf}X)_i := \bar{\nabla}_j(\mathcal{L}_{conf,\bar{g}}X)^j_i = f_i \tag{2.27}$$

These equations have one and only one solution $(\lambda, X) \in W^p_{2,\delta} \times W^p_{2,\delta}$ if $p > \frac{n}{2}$, $-\frac{n}{p} < \delta < n - 2 - \frac{n}{p}$ (see Appendix II). Then every pair (h, k) has a unique continuous decomposition in $W^p_{2,\delta} \times W^p_{1,\delta+1}$

$$(h, k) = (h_1, k_1) + (h_2, k_2) \tag{2.28}$$

with $(h_1, k_1) \in E_1 := W^p_{2,\delta} \cap \{\delta\Phi = 0\}$ and (h_2, k_2) of the form (2.24). Therefore the kernel E_1 has a closed complement E_2, the space of such pairs.

For the proof in the case $trK \not\equiv 0$, see the quoted article of Corvino and Schoen. This linearization stability is also a consequence of the theorem of existence proved in Section (12.2).

3 CF (Conformally Formulated) constraints

In the conformal method one specifies an n-dimensional manifold and a non-physical Riemannian metric $\tilde{g}$, often in the following denoted γ, on M. One looks for a physical space metric $\bar{g}$ conformal to γ. The Hamiltonian constraint reads then as a semilinear elliptic equation – now called the Lichnerowicz equation – for the conformal factor, with principal operator the Laplacian in the metric γ and coefficients depending on K and on the sources. The extrinsic curvature K on M is also split into its specified trace τ, the mean curvature of M in the ambient spacetime, and a non-physical traceless symmetric 2-tensor $\tilde{K}$ weighted by a well-chosen power of the conformal factor of the metric. The momentum constraint reads then as a first-order linear system for $\tilde{K}$ when the metric γ, the conformal factor and the sources are known[13]. This system is turned into an elliptic system[14] for a vector field X by using the expression of $\tilde{K}$ as the sum of the Lie derivative of X and a specified tensor. York[15] wrote another elliptic system for the momentum constraint using the splitting of symmetric traceless 2-tensors into a transverse (divergence-free) part and a conformal Lie derivative. These splittings lead to an elliptic system for the coupled constraints, with coefficients depending on specified data.

The conformal method transforms the constraints into a semilinear elliptic system. A remarkable property is that in vacuum the momentum constraint is independent of the conformal factor if τ is a constant ($\tau = 0$; Lichnerowicz (1944), general case York (1972)). In this case, to solve the constraints, one solves first a linear elliptic system, and then inserts its solution in the Lichnerowicz equation, the only non-linearity remaining in the problem. A scaling of the sources which permits the extension of this decoupling property to the non-vacuum case has been introduced by York (1972). One replaces in the equations the physical

[13] Lichnerowicz, A. (1944) *J. Math. Pures Appl.*, **23**, 39–63. He used this system to construct on R^3, in the time symmetric case ($K = 0$) conformally flat initial data for the n-body problem with unscaled masses and zero momentum.

[14] Choquet-Bruhat, Y. (1971) *Commun. Math. Phys.*, **21**, 211–18.

[15] York, J. W. (1972) *Phys. Rev. Lett.*, **28**, 1082 and York, J. W. (1974) *Ann. Inst. Henri Poincaré*, **21**, 319–32.

quantities by the product of specified tilde quantities multiplied by a power of the conformal factor. Since the tilde quantities cannot be observed there is a large arbitrariness in the choice of their scaling, though some justifications are given *a posteriori*.

We remark that, in fact, the physical quantities themselves, involved in astronomical or cosmological problems, are not directly observable. The tilde quantities play the role of parameters to construct initial data solutions of the constraints which can, hopefully, be used for evolution and lead to results which can be confronted with observations.

We treat the case of general n, specifying other values only when they enjoy special properties.

3.1 *Hamiltonian constraint*

In order to turn the Hamiltonian constraint into an elliptic equation for a scalar function, one considers the metric $\bar{g}$ as given up to a conformal factor; one sets

$$\bar{g} = e^{2\lambda}\gamma, \quad \text{i.e.} \quad \bar{g}_{ij} = e^{2\lambda}\gamma_{ij}$$

with γ a given metric on M and λ a function to be determined.

The scalar curvatures of the conformal metrics $\bar{g}$ and γ are linked by the formula (see Appendix VI)

$$\bar{R} \equiv e^{-2\lambda}\{R(\gamma) - 2(n-1)\Delta_\gamma\lambda - (n-1)(n-2)\bar{g}^{ij}\partial_i\lambda\partial_j\lambda\}$$

When $n = 1$, $\bar{R}$ is proportional to $R(\gamma)$. When $n = 2$ the quantity between brackets above is a linear second-order operator in λ. When $n \geq 3$, we set

$$e^{2\lambda} = \varphi^{2p}, \quad \text{i.e.} \quad \bar{g}_{ij} = \varphi^{2p}\gamma_{ij},$$

and we choose p in such a way that the operator on φ appearing within the brackets is linear in λ; this goal is attained by choosing

$$p = \frac{2}{n-2}, \quad \text{i.e.} \quad \bar{g}_{ij} = \varphi^{\frac{4}{n-2}}\gamma_{ij}, \quad \bar{g}^{ij} = \varphi^{-\frac{4}{n-2}}\gamma^{ij},$$

which we suppose from now on. The following identity then holds (as may be seen by a straightforward calculation):

$$R(\bar{g}) \equiv \varphi^{-(n+2)/(n-2)}\left(\varphi R(\gamma) - \frac{4(n-1)}{n-2}\Delta_\gamma\varphi\right), \qquad \bar{g} \equiv \varphi^{4/(n-2)}\gamma. \tag{3.1}$$

The Hamiltonian constraint becomes a semilinear elliptic equation for φ, with a non-linearity of a fairly simple type, when γ, K, and ρ are known, namely

$$\Delta_\gamma\varphi - k_n R(\gamma)\varphi + k_n(|K|^2_{\bar{g}} - \tau^2 + 2\rho)\varphi^{(n+2)/(n-2)} = 0, \tag{3.2}$$

with

$$k_n = \frac{n-2}{4(n-1)} \quad \text{and} \quad \tau := tr_{\bar{g}}K, \quad \text{i.e.} \quad \tau := \bar{g}^{ij}K_{ij}\ .$$

3.2 *Momentum constraint*

We can express the momentum constraint in terms of γ, φ, K, and J by using the relations between the connections of two conformal metrics.

Lemma 3.1 *On an n-dimensional manifold, if $\bar{g} = \varphi^{4/(n-2)}\gamma$, the covariant derivatives in $\bar{g}$ and γ being respectively denoted $\bar{\nabla}$ and D, the divergences in the metric $\bar{g}$ and γ of an arbitrary contravariant 2-tensor P^{ij} are linked by the identity*

$$\bar{\nabla}_i P^{ij} \equiv \varphi^{-2(n+2)/(n-2)} D_i\{\varphi^{2(n+2)/(n-2)} P^{ij}\} - \frac{2}{n-2}\varphi^{-1}\gamma^{ij}\partial_i\varphi tr_\gamma P. \quad (3.3)$$

Proof The proof is by simple computation using the identity which links the coefficients of the connections $\bar{\Gamma}$ and C of $\bar{g}$ and γ,

$$\bar{\Gamma}^i_{jh} = C^i_{jh} + \frac{2}{n-2}\varphi^{-1}\{\delta^i_j\partial_h\varphi + \delta^i_h\partial_j\varphi - \gamma^{ik}\gamma_{jh}\partial_k\varphi\}. \quad (3.4)$$

□

One sees from the identity (3.3) that it is convenient to split the unknown K into a weighted traceless part and its trace τ, namely we set

$$K^{ij} = \varphi^{-2(n+2)/(n-2)}\tilde{K}^{ij} + \frac{1}{n}\bar{g}^{ij}\tau. \quad (3.5)$$

Equivalently, indices in K and $\tilde{K}$ being respectively lowered with $\bar{g}$ and γ

$$K_{ij} = \varphi^{-2}\tilde{K}_{ij} + \frac{1}{n}\bar{g}_{ij}\tau, \quad \bar{g}_{ij} = \varphi^{\frac{4}{n-2}}\gamma_{ij}. \quad (3.6)$$

The tensor $\tilde{K}$ is symmetric and traceless, that is:

$$tr_\gamma\tilde{K} \equiv \gamma^{ij}\tilde{K}_{ij} = \varphi^{\frac{2n}{n-2}}\bar{g}^{ij}\left(K_{ij} - \frac{1}{n}\bar{g}_{ij}\tau\right) = 0.$$

The momentum constraint (1.2) becomes

$$D_i\tilde{K}^{ij} = \frac{n-1}{n}\varphi^{2n/(n-2)}\gamma^{ij}\partial_i\tau + \varphi^{2(n+2)/(n-2)}J. \quad (3.7)$$

This equation has the interesting property not to contain φ when J is zero and τ is constant on M, that is when M is a submanifold with constant mean curvature in the ambient spacetime. In this case the momentum constraint reduces to the linear homogeneous system for $\tilde{K}$:

$$D_j\tilde{K}^{ij} = 0, \quad \gamma_{ij}\tilde{K}^{ij} = 0. \quad (3.8)$$

Symmetric 2-tensors satisfying (3.8) are called **TT tensors** (Transverse, Traceless). Calculus using Lemma 3.1 shows that the space of TT tensors is the same for two conformally related metrics.

The definition (3.5) gives

$$|K|^2_{\bar{g}} := \bar{g}_{ih}\bar{g}_{jk}K^{ij}K^{hk} = \varphi^{(-3n+2)/(n-2)}\gamma_{ih}\gamma_{jk}\tilde{K}^{ij}\tilde{K}^{hk} + \frac{1}{n}\tau^2$$
$$\equiv \varphi^{(-3n+2)/(n-2)}|\tilde{K}|^2_{\gamma} + \frac{1}{n}\tau^2 \tag{3.9}$$

The Hamiltonian constraint reads therefore

$$\Delta_\gamma\varphi - k_n R(\gamma)\varphi + k_n\varphi^{(-3n+2)/(n-2)}|\tilde{K}|^2_{\gamma} - \frac{n-2}{4n}\varphi^{(n+2)/(n-2)}\tau^2$$
$$+ \frac{n-2}{2(n-1)}\rho\varphi^{(n+2)/(n-2)} = 0 \tag{3.10}$$

It is a semilinear elliptic equation for φ when γ, a Riemannian metric, τ, ρ, and $\tilde{K}$ are known. It is called the **Lichnerowicz equation**[16].

3.3 *Scaling of the sources*

3.3.1 *Decoupling theorem*

Suppose one is given a metric γ which will be conformal to the physical metric $\bar{g}$. The expression $\varphi^{\frac{4}{n-2}}$ of the conformal factor is just a formal choice. The problem is to find a representation of the initial physical scalar energy density and the vector momentum density

$$\bar{\rho} := \bar{N}^{-2}\bar{T}_{00}, \quad \bar{J}^i := -\bar{N}^{-1}\bar{T}^i_0 \tag{3.11}$$

as products of power of φ with quantities which it is physically plausible to suppose to be known. The choice is somewhat arbitrary. The usual choice relies on heuristic arguments which depend on the nature of the source fields, and has the advantage of leading to tractable mathematics. Generalizing the definition given by York in the case $n = 3$ we say that a source is **York-scaled** if it holds that (we overline physical quantities on the initial manifold $(M, \bar{g})$, and put a tilde to specified quantities):

$$\bar{J} = \varphi^{-2(n+2)/(n-2)}\tilde{J}\ , \quad \bar{\rho} = \tilde{\rho}\varphi^{-2(n+1)/(n-2)} \tag{3.12}$$

where $\tilde{J}$ and $\tilde{\rho}$ are specified on the initial conformal manifold (M, γ).

A fundamental result is the following theorem, whose proof can be read directly from the momentum Equation (3.7).

Theorem 3.2 *(York) When the momentum is York-scaled the momentum constraint decouples from the Lichnerowicz equation if the mean curvature τ is a constant.*

[16] It was obtained by Lichnerowicz (1944) for $n = 3$, extended to general n by CB 1996. Scaling of the sources was introduced by York (1972). We will still call Lichnerowicz equations all the equations deduced from the Hamiltonian constraint by the conformal method.

It seems reasonable to keep the "tilde" quantities equal to the physical quantities when these are independent of the choice of γ and also of the embedding of M in spacetime; that is, of the choice of a spacetime observer. Otherwise, the choice is more ambiguous. A possible procedure is to introduce the densitized lapse[17]. Then one defines the unphysical lapse $\tilde{N}$ associated with the metric γ as being the scalar such that $\tilde{N}$ and $\bar{N}$, the initial value of the physical lapse, have the same associated densitized lapse respectively for γ and $\bar{g}$, that is:

$$\bar{N}(Det\bar{g})^{-\frac{1}{2}} = \tilde{N}(Det\gamma)^{-\frac{1}{2}} \tag{3.13}$$

i.e.

$$\bar{N} = \varphi^{2n/(n-2)}\tilde{N}. \tag{3.14}$$

Remark that in the presence of sources of different types in the model under study, only one type of specified lapse can be introduced, but one can define *ad hoc* scalings without lapse choice justification.

We treat the fundamental examples in the following section.

3.3.2 *Source a 2-form F*

The stress energy tensor (indices raised with the physical spacetime metric) is

$$T_{\alpha\beta} \equiv F_{\alpha}{}^{\lambda}F_{\beta\lambda} - \frac{1}{4}g_{\alpha\beta}F_{\lambda\mu}F^{\lambda\mu}. \tag{3.15}$$

The physical momentum density on spacetime is:

$$J^i = -N^{-1}T_0^i = -N^{-1}F_0{}^jF^i{}_j. \tag{3.16}$$

The physical "electric" part of F with respect to a physical observer at rest on the initial manifold M, $x^0 = constant$, is the vector field on M given by the contraction of F with the unit normal, that is in the Cauchy adapted frame:

$$\bar{E}^i =: \bar{N}^{-1}\bar{F}_0{}^i. \tag{3.17}$$

The physical "magnetic" part of F is the 2-form $\bar{F}$ on the initial manifold M with components $\bar{F}_{ij}$, supposed to be known. If, inspired by the densitized lapse hypothesis, one gives as conformal data the vector $\tilde{E}$ and the 2-form $\tilde{F}$ such that

$$\tilde{E}^i := \tilde{N}^{-1}\bar{F}_0{}^i \equiv \varphi^{\frac{2n}{n-2}}\bar{E}^i, \quad \tilde{F}_{ij} = \bar{F}_{ij}, \tag{3.18}$$

then:

$$\bar{J}^i = \varphi^{\frac{-2(n+2)}{n-2}}\tilde{J}^i, \quad \text{with} \quad \tilde{J}^i \equiv \tilde{E}^i\gamma^{ik}\tilde{F}_{kj}, \tag{3.19}$$

which shows that the momentum is York-scaled for any n.

[17] Used in the Choquet-Bruhat and Ruggeri (1983) hyperbolic fomulation, in Teitelboim's (1983) Hamiltonian study (*Phys. Rev. D*, **28**, 297–311, and previously by others in various contexts, in particular by Taub.

On the other hand it holds that:

$$\bar{\rho} \equiv \bar{N}^{-2}\bar{T}_{00} = \frac{1}{2}\bar{N}^{-2}\bar{F}_{0i}\bar{F}_0{}^i + \frac{1}{4}\bar{F}^{ij}\bar{F}_{ij}. \tag{3.20}$$

It results from an elementary computation that:

$$\bar{N}^{-1}\bar{F}_{0i} \equiv \bar{N}^{-1}\bar{g}_{ij}\bar{F}_0{}^i = \varphi^{-2}\gamma_{ij}\tilde{E}^i, \tag{3.21}$$

hence the following "electric part",

$$\bar{N}^{-2}\bar{F}_{0i}\bar{F}_0{}^i = \varphi^{-\frac{2(n+1)}{n-2}}\gamma_{ij}\tilde{E}^i\tilde{E}^j, \tag{3.22}$$

is York-scaled; but the magnetic part,

$$\bar{F}^{ij}\bar{F}_{ij} = \varphi^{-\frac{8}{n-2}}\gamma^{ih}\gamma^{jk}\tilde{F}_{ij}\tilde{F}_{hk}, \tag{3.23}$$

is York-scaled only in the case $n = 3$.

The first Maxwell constraint, $d\bar{F} = 0$, is metric-independent. The second Maxwell constraint reduces to the space divergence of the electric field E. Indeed, it reads in vacuum

$$\overline{\nabla_\alpha F_0{}^\alpha} \equiv \overline{\partial_\alpha F_0{}^\alpha - \omega^\beta_{\alpha 0}F_\beta{}^\alpha + \omega^\alpha_{\alpha\beta}F_0{}^\beta} = 0.$$

We find by a straightforward calculation, using the definition of $\bar{E}$ and the formulae in Chapter 6 for the connection ω on spacetime and the antisymmetry of F (recall that ω^0_{ij} is symmetric), that this constraint can be written:

$$\overline{\nabla_\alpha F_0{}^\alpha} \equiv \bar{N}\bar{\nabla}\bar{E}^i = 0. \tag{3.24}$$

It is trivial to check, using the formula for the divergence of a vector field, that

$$\bar{\nabla}_i\bar{E}^i \equiv \frac{1}{\sqrt{\det\bar{g}}}\partial_i(\sqrt{\det\bar{g}}\bar{g}^{ij}\partial_j\bar{E}^i) = \varphi^{\frac{-2n}{n-2}}\tilde{\nabla}_i\tilde{E}^i. \tag{3.25}$$

Therefore the second Maxwell constraint, in the absence of electric charges, is conformally invariant. We have shown that:

In the conformal formulation with a densitized lapse choice, the Maxwell constraints and the Einstein momentum constraint decouple from the Hamiltonian constraint.

Case $n = 3$.

One associates with the 2-form $\bar{F}_{ij}$ its adjoints, respectively in the metrics $\bar{g}$ and γ, which define the magnetic vector fields $\bar{H}$ and $\tilde{H}$. These are linked by the relation

$$\bar{H}^i = \frac{1}{2}\bar{\eta}^{ijk}\bar{F}_{jk} = \frac{1}{2}\varphi^{-6}\tilde{\eta}^{ijk}\tilde{F}_{jk} \equiv \varphi^{-6}\tilde{H}^i \tag{3.26}$$

with $\tilde{\eta}$ and $\bar{\eta}$ respectively the volume forms of γ and $\bar{g}$.

We consider that $\tilde{H}$ is the known magnetic vector, and we verify that the momentum source is properly scaled for the decoupling when $\partial\tau = 0$:

$$\bar{J}^i = \varphi^{-10}\tilde{J}^i, \quad \text{with} \quad \tilde{J}^i = -\gamma^{ik}\tilde{\eta}_{kjl}\tilde{E}^j\tilde{H}^l\ . \tag{3.27}$$

The energy density is

$$\bar{\rho} = \frac{1}{2}\bar{g}_{ij}(\bar{E}^i\bar{E}^j + \bar{H}^i\bar{H}^j) \equiv \varphi^{-8}\tilde{\rho} \tag{3.28}$$

with $\tilde{\rho}$, considered as known on M, given by

$$\tilde{\rho} \equiv \frac{1}{2}\gamma_{ij}(\tilde{E}^i\tilde{E}^j + \tilde{H}^i\tilde{H}^j). \tag{3.29}$$

We find again that, when $n = 3$, the sources are scaled as defined by York.

Analogous results hold when F takes its values in a given normed vector space, for instance when F is a **Yang–Mills** field.

3.3.3 *Case of a scalar field*

The energy density on M of a scalar field ψ with potential $V(\psi)$, for an observer at rest in the physical frame is

$$\bar{\rho} = \frac{1}{2}(\bar{\pi}^2 + \bar{g}^{ij}\partial_i\bar{\psi}\partial_j\bar{\psi}) + V(\bar{\psi}), \quad \text{with} \quad \bar{\pi} := \bar{N}^{-1}\overline{\partial_0\psi} \tag{3.30}$$

The values $\bar{\psi}$, $\partial_i\bar{\psi}$, and $\overline{\partial_0\psi}$ of ψ, $\partial_i\psi$, and $\partial_0\psi$ on M are independent of the choice of γ and N; but it holds if $\tilde{N}$ is a densitized lapse, that

$$\bar{\pi} \equiv \bar{N}^{-1}\overline{\partial_0\psi} = \varphi^{-2n/(n-2)}\tilde{\pi} \quad \text{where} \quad \tilde{\pi} := \tilde{N}^{-1}\overline{\partial_0\psi}.$$

Then $\bar{\rho}$ reads:

$$\bar{\rho} = \frac{1}{2}(\varphi^{\frac{-4n}{n-2}}|\tilde{\pi}|^2 + \varphi^{\frac{-4}{n-2}}\gamma^{ij}\partial_i\bar{\psi}\partial_j\bar{\psi}) + V(\bar{\psi}). \tag{3.31}$$

Neither of the terms in $\bar{\rho}$ are York-scaled. The term with $|\tilde{\pi}|^2$ adds in the Hamiltonian constraint to $|\tilde{K}|^2_\gamma$, while the term $V(\bar{\psi})$ remains unscaled by a power of φ. The middle term gives a positive contribution to the φ term which adds to $-R(\gamma)$.

The momentum density is

$$\bar{J}^i = -\bar{N}^{-1}\bar{g}^{ij}\partial_j\bar{\psi}\overline{\partial_0\psi} = -\varphi^{-2(n+2)/(n-2)}\gamma^{ij}\partial_j\bar{\psi}\tilde{\pi} = \varphi^{-2(n+2)/(n-2)}\tilde{J}^i. \tag{3.32}$$

We see that the momentum is York-scaled: *the Einstein scalar constraints decouple if* $\partial\tau = 0$.

A particular case of a scalar field is a **cosmological constant** Λ. Then $\partial_\alpha\psi$ has to be replaced by 0, and $V(\psi)$ by Λ.

We state the following theorem, which is a consequence of the York theorem.

Theorem 3.3 *With a densitized lapse scaling, the momentum and Hamiltonian constraints decouple if the sources are only Maxwell (or Yang–Mills) and scalar fields.*

3.3.4 *Matter sources*

We first show that the constraints with matter sources decouple with an appropriate choice of the specified quantities, which is not interpreted as densitized lapse.

Consider the case of a perfect fluid

$$T_{\alpha\beta} \equiv (\mu + p)u_\alpha u_\beta + p g_{\alpha\beta} \tag{3.33}$$

with μ and p scalars and u a unit vector in the physical metric; that is, its components in the Cauchy frame $\theta^0 = dt$, $\theta^i = dx^i + \beta^i dt$, at a point of the initial manifold, are such that:

$$\bar{N}^2(\bar{u}^0)^2 - \bar{g}_{ij}\bar{u}^i\bar{u}^j = 1. \tag{3.34}$$

The terms $\bar{\rho}$ and $\bar{J}$ are

$$\bar{\rho} := (\bar{N}\bar{u}^0)^2(\bar{\mu} + \bar{p}) - \bar{p}, \quad \bar{J}^i := \bar{N}(\bar{\mu} + \bar{p})\bar{u}^0\bar{u}^i \tag{3.35}$$

The components of a specified vector $\tilde{u}$, in the same direction as u, which is unit in the specified conformal metric γ and some specified lapse $\tilde{N}$ satisfies the equation

$$\tilde{N}^2(\tilde{u}^0)^2 - \gamma_{ij}\tilde{u}^i\tilde{u}^j = 1. \tag{3.36}$$

The conformal relation between $\bar{g}$ and γ shows that:

$$\tilde{u}^i = \varphi^{\frac{2}{n-2}}\bar{u}^i \tag{3.37}$$

while, for corresponding choices of $\tilde{N}$ and $\tilde{u}^0$, it must hold that

$$\bar{N}\bar{u}^0 = \tilde{N}\tilde{u}^0 = (\gamma_{ij}\tilde{u}^i\tilde{u}^j + 1)^{\frac{1}{2}}. \tag{3.38}$$

The momentum $\bar{J}$ reads then

$$\bar{J}^i := \bar{N}\bar{u}^0(\bar{\mu} + \bar{p})\bar{u}^i = \varphi^{-\frac{2}{n-2}}\tilde{N}\tilde{u}^0\tilde{u}^i(\bar{\mu} + \bar{p}). \tag{3.39}$$

This momentum is York-scaled if $\bar{J}^i = \varphi^{-\frac{2(n+2)}{n-2}}\tilde{J}^i$; that is, if $\tilde{\mu}$ and $\tilde{p}$ are such that

$$\bar{\mu} = \varphi^{\frac{-2(n+1)}{n-2}}\tilde{\mu}, \quad \bar{p} = \varphi^{\frac{-2(n+1)}{n-2}}\tilde{p}, \quad \text{then} \quad \tilde{J}^i \equiv (\gamma_{ij}\tilde{u}^i\tilde{u}^j + 1)^{\frac{1}{2}}\tilde{u}^i(\tilde{\mu} + \tilde{p}). \tag{3.40}$$

With such a scaling the source appearing in the Hamiltonian constraint is

$$\bar{\rho} = \varphi^{\frac{-2(n+1)}{n-2}}[(\tilde{\mu} + \tilde{p})(\tilde{N}\tilde{u}^0)^2 - \tilde{p}] \equiv \varphi^{\frac{-2(n+1)}{n-2}}\tilde{\rho}. \tag{3.41}$$

It is also York-scaled.

A possible interpretation of the considered $\tilde{\mu}$ and $\tilde{p}$ in relation with physical quantities comes from the equality of spacetime matter densities

$$\bar{\mu}\bar{N}(det\bar{g})^{\frac{1}{2}} = \tilde{\mu}\tilde{N}(det\gamma)^{\frac{1}{2}} \tag{3.42}$$

when one supposes that $\bar{N}$ is conformally scaled by the same factor as $\bar{g}$. This is not the densitized lapse choice.

3.3.5 *Unscaled matter fields*

Since $\bar{\mu}$ and $\bar{p}$ are scalar quantities one is tempted not to scale them; that is to keep them as specified quantities. We cannot independently specify $\bar{u}^i$, $\bar{u}^0$, and $\bar{N}$, since u is a unit vector in the physical spacetime metric. To obtain polynomial equations in φ, we must still scale u by (3.37) and (3.38). Then in the Hamiltonian constraint the term $\bar{\rho}$ adds to the τ^2 term and to the potential of the scalar field. The momentum $\bar{J}$ reads, with $\tilde{J}$ a specified vector

$$\bar{J}^i = \bar{N}\bar{u}^0\bar{\rho}\varphi^{-\frac{2}{n-2}}\tilde{u}^i := \varphi^{-\frac{2}{n-2}}\tilde{J}^i \tag{3.43}$$

The constraints do not decouple if $\tilde{J} \neq 0$: a term $\varphi^{\frac{2(n+1)}{n-2}}\tilde{J}$ appears in the momentum constraint.

3.4 *Summary of results*

We summarize the previous results when the sources are of the type listed in the previous subsection. A Riemannian metric γ is arbitrarily chosen on the manifold M, together with a scalar function τ. The sources are supposed to be known, up to some conformal weight depending on their nature and the choice of their scaling, possibly linked with the choice of the scaling of the lapse $\bar{N}$ on M.

The **Hamiltonian constraint** is always equivalent to a Lichnerowicz equation

$$\mathcal{H} \equiv \Delta_\gamma\varphi - f(.,\varphi) = 0, \tag{3.44}$$

and the **momentum constraint** to the system

$$\mathcal{M}^i \equiv D_j\tilde{K}^{ij} - F^i = 0. \tag{3.45}$$

$$F^i := \frac{n-1}{n}\varphi^{2n/(n-2)}\gamma^{ij}\partial_j\tau + \bar{J}^i\varphi^{2(n+2)/(n-2)}$$

In vacuum $\bar{J} \equiv 0$ and, with $k_n = \frac{n-2}{4(n-1)}$

$$f(.,\varphi) \equiv k_n R(\gamma)\varphi - k_n|\tilde{K}|^2_\gamma\varphi^{(-3n+2)/(n-2)} + \frac{n-2}{4n}\tau^2\varphi^{(n+2)/(n-2)} \tag{3.46}$$

In the presence of sources, the CF constraints depend on a choice of the scaling of the sources and of the lapse.

In all cases, the formula (3.6) gives the solution $(\bar{g}, K)$ of the constraints corresponding to the solution $(\varphi, \tilde{K})$ of the CF constraints with geometric data γ, τ, and the conformal source data.

A fairly general system containing the usual fields and matter is obtained by the following equalities.

The function f in the *Lichnerowicz equation* is

$$f(.,\varphi) \equiv r\varphi - a\varphi^{-\frac{3n-2}{n-2}} - q_1\varphi^{-\frac{n}{n-2}} - q_2\varphi^{\frac{-6+n}{n-2}} + b\varphi^{\frac{n+2}{n-2}}, \tag{3.47}$$

$$r \equiv k_n(R(\gamma) - |D\bar{\psi}|^2_\gamma), \quad a \equiv k_n(|\tilde{K}|^2_\gamma + |\tilde{\pi}|^2), \quad b \equiv \frac{n-2}{4n}\tau^2 - q_0.$$

It holds that $a \geq 0$.

The coefficients q_1 and q_2 correspond to the electric and magnetic energies with the densitized lapse scaling, and also to matter sources scaled as in (3.34). They are given by:

$$q_1 = k_n(\gamma_{ij}\tilde{E}^i\tilde{E}^j + 2\tilde{\rho}), \qquad q_2 = k_n\gamma^{ih}\gamma^{jk}\bar{F}_{ij}\bar{F}_{hk}, \tag{3.48}$$

hence $q_1 \geq 0$, $q_2 \geq 0$. Remark that they are multiplied by the same power φ^{-3} of φ only if $n = 3$. Moreover the sign of the power of φ with coefficient q_2 is negative only if $n < 6$. This would play a role in the solution of constraints when $n \geq 6$ and non-zero magnetic field.

The coefficient q_0 corresponds to non-scaled energies. If there is a scalar field with potential $V(\psi)$ and unscaled matter sources with matter density $\bar{\rho}_0$, then

$$q_0 = 2k_n(V(\bar{\psi}) + \bar{\rho}_0). \tag{3.49}$$

The sign of q_0 is not necessarily a constant on M.

The source in the *momentum constraint* is

$$\bar{J}^i\varphi^{2(n+2)/(n-2)} \equiv \tilde{J}_1^i + \varphi^{2(n+1)/(n-2)}\tilde{J}_0^i. \tag{3.50}$$

The φ-unweighted courant $\tilde{J}_1$ comes from the scalar and electromagnetic fields and scaled matter sources; it is given by

$$\tilde{J}_1^i \equiv -\gamma^{ij}(\partial_j\bar{\psi}\tilde{\pi} + \tilde{\eta}_{jkl}\tilde{E}^k\tilde{H}^l) + (1 + \gamma_{hj}\tilde{u}^h\tilde{u}^j)^{\frac{1}{2}}\tilde{u}^i(\tilde{\mu} + \tilde{p}) \tag{3.51}$$

The φ-weighted courant $\tilde{J}_0$ comes from unscaled matter source. It is

$$\tilde{J}_0^i \equiv (1 + \gamma_{hj}\tilde{u}^h\tilde{u}^j)^{\frac{1}{2}}\tilde{u}^i(\bar{\mu} + \bar{p}). \tag{3.52}$$

This current $\tilde{J}_0$ vanishes if $\tilde{u}^i$ vanishes on the support of unscaled matter sources. If $\tilde{J}_0$ vanishes, the momentum constraint does not contain φ if $\partial_i\tau \equiv 0$, i.e. if the initial manifold has constant mean extrinsic curvature.

3.5 *Conformal transformation of the CF constraints*

It is intuitive that, in some sense, the solution of Einstein's constraints should not depend on the choice of the specified conformal metric in a conformal class; that is, the two systems of CF constraints obtained from two different choices γ and γ', together with appropriate rescaling in the presence of sources, should be equivalent.

Theorem 3.4 *Suppose that $(\tilde{K}^{ij}, \varphi)$ is a solution of the vacuum CF constraints for the metric γ and a given funtion τ. Then the CF vacuum constraints in the metric $\gamma'_{ij} = \theta^{4/(n-2)}\gamma_{ij}$, with $\tau' = \tau$, admit the solution $\tilde{K}'^{ij} = \theta^{-2(n+2)/(n-2)}\tilde{K}^{ij}$, $\varphi' = \theta^{-1}\varphi$.*

The same result holds in the presence of sources if the sources q_0, q_1, q_2, $\tilde{J}_0$, $\tilde{J}_1$ are determined from the physical data by appropriate rescaling.

Proof We set:

$$\gamma'_{ij} \equiv \theta^{4/(n-2)}\gamma_{ij}, \tag{3.53}$$

then

$$|D\bar{\psi}|^2_{\gamma} \equiv \theta^{4/(n-2)}|D\bar{\psi}|^2_{\gamma'} \tag{3.54}$$

This identity together with the following one, a generalization of (3.1), which can be found in Appendix VI,

$$\Delta_\gamma\varphi - k_n R(\gamma)\varphi \equiv \theta^{(n+2)/(n-2)}\{\Delta_{\gamma'}\varphi' - k_n R(\gamma')\varphi'\}, \quad \varphi' \equiv \theta^{-1}\varphi, \tag{3.55}$$

show that, with r given in (3.50)

$$\Delta_\gamma\varphi - r\varphi \equiv \theta^{(n+2)/(n-2)}\{\Delta_{\gamma'}\varphi' - r'\varphi'\}, \quad \varphi' \equiv \theta^{-1}\varphi. \tag{3.56}$$

On the other hand the application of Lemma 3.1 to the metrics γ' and γ and a symmetric traceless tensor gives

$$D'_i\tilde{K}'^{ij} \equiv \theta^{-2(n+2)/(n-2)}D_i\tilde{K}^{ij}, \quad \tilde{K}^{ij} \equiv \theta^{2(n+2)/(n-2)}\tilde{K}'^{ij}. \tag{3.57}$$

The definition of q_2 gives

$$q'_2 = \theta^{\frac{-8}{n-2}}q_2. \tag{3.58}$$

The densitized lapse choice – which implies $\tilde{N}' = \theta^{2n/(n-2)}\tilde{N}$ – for the scaling of the time derivative of the scalar field leads to

$$\tilde{\pi}' \equiv \theta^{-2n/(n-2)}\pi. \tag{3.59}$$

Therefore the coefficients a and a' are linked by

$$a' \equiv k_n(\tilde{K}'.'\tilde{K}' + (\tilde{\pi}')^2) \equiv \theta^{-4n/(n-2)}a. \tag{3.60}$$

If we assume also the densitized lapse choice for scaling of the electromagnetic field, but suppose that the specified scaled matters are linked by the relation

$$\tilde{\rho}' = \theta^{-4/(n-2)}\tilde{\rho}, \tag{3.61}$$

then

$$q'_1 = \theta^{-4/(n-2)}q_1. \tag{3.62}$$

Choosing $\tau' = \tau$, and $\bar{\rho}' = \bar{\rho}$ for unscaled matter, we find that

$$b' \equiv b \quad \text{and} \quad \gamma'^{ij}\partial_i\tau = \theta^{\frac{-4}{n-2}}\gamma^{ij}\partial_i\tau. \tag{3.63}$$

We have then, using $\varphi' = \theta^{-1}\varphi$,

$$\mathcal{H}'(\tilde{K}', \varphi') \equiv \theta^{-(n+2)/(n-2)}\mathcal{H}(K, \varphi). \tag{3.64}$$

Analogous arguments give

$$\mathcal{M}'^i(\tilde{K}', \varphi') \equiv \theta^{-2(n+2)/(n-2)}\mathcal{M}^i(K, \varphi) \tag{3.65}$$

□

3.6 *The momentum constraint as an elliptic system*

3.6.1 *York splitting*

As is the case for general linear systems, an arbitrary solution of the momentum constraint (3.45) is the sum of a solution of the associated homogeneous system, i.e. here an arbitrary TT tensor $\tilde{K}_{TT}$, and a particular solution. Usually one looks for a particular solution in the formal L^2 dual of the kernel of the space of solutions of the homogeneous system, here the space of TT tensors.

Lemma 3.5 *(**York splitting**) The formal L^2 dual of the space of TT tensors in a metric γ is the space of the conformal Lie derivative of γ with respect to some vector field Y, that is of the space of symmetric traceless 2-tensors of the form*

$$(\mathcal{L}_{\gamma,conf}Y)^{ij} \equiv D^iY^j + D^jY^i - \frac{2}{n}\gamma^{ij}D_kY^k. \tag{3.66}$$

If a tensor of the form (3.66) satisfies the momentum constraint (3.45), then it satisfies the second-order system

$$\Delta_{\gamma,conf}Y := D.(\mathcal{L}_{\gamma,conf}Y) = F.$$

This system is shown in the appendix II "Elliptic systems" to be elliptic.

The search for the arbitrary TT tensor $\tilde{K}^{ij}_{TT}$ can itself be done through the data of an arbitrary traceless tensor U^{ij} by solving for a vector Z the following elliptic system

$$\Delta_{\gamma,conf}Z \equiv D.(\mathcal{L}_{\gamma,conf}Z) = -D.U,$$

the traceless tensor $\mathcal{L}_{\gamma,conf}Z - U$ is then also transverse. The arbitrary data in the extrinsic curvature K is in this scheme the symmetric traceless tensor U^{ij}. The physical extrinsic curvature is given by

$$K^{ij} = \varphi^{-2(n+2)/(n-2)}[(\mathcal{L}_{\gamma,conf}X)^{ij} + U^{ij}] + \frac{1}{n}\bar{g}^{ij}\tau,$$

where the vector field $X := Y + Z$ is a solution of the system, which is linear and elliptic when F is assumed to be known,

$$\Delta_{\gamma,conf}X = F, \tag{3.67}$$

with now

$$F := -D.U + \frac{n-1}{n}\varphi^{2n/(n-2)}\partial\tau + \varphi^{2(n+2)/(n-2)}J$$

with, in the scalings given in Section 3.3,

$$\varphi^{2(n+2)/(n-2)}J = \tilde{J}_1 + \tilde{J}_0\varphi^{4/(n-2)}.$$

Remark 3.6 *The spaces of conformal Killing vector fields (these spaces are the kernels of the conformal Killing operators) of (M,γ) and $(M,\bar{g})$ are the same; use the definition (3.66) to check it.*

3.6.2 *Conformally invariant (thin sandwich) splitting*

It was remarked by York in 1999 that the splitting of the solution $\tilde{K}$ into a given traceless tensor U and the conformal Lie derivative of an unkown vector X cannot be made conformally invariant. To have $\tilde{K}' \equiv \theta^{-2(n+2)/(n-2)}\tilde{K}$ when $\gamma' = \theta^{\frac{4}{n-2}}\gamma$ we can impose the relation between the given traceless tensors U and U':

$$U'^{ij} \equiv \theta^{-2(n+2)/(n-2)}U^{ij}, \tag{3.68}$$

but for an arbitrary vector X it holds that

$$(\mathcal{L}_{\gamma',conf}X)^{ij} \equiv \theta^{-4/(n-2)}(\mathcal{L}_{\gamma,conf}X)^{ij}. \tag{3.69}$$

No scaling of X by a power of φ leads to a vector X' and just another scaling of its conformal Lie derivative. This lack of conformal invariance makes difficult the physical interpretation of the given tensor U since the splitting into known U and unknown X depends on the choice of γ, and not only on its conformal class. York proposed to remedy this defect by what he called "the **conformal thin sandwich** formulation". Inspired by his work, and the expression $K^{ij} \equiv N^{-1}\bar{\partial}_0 g^{ij}$, we replace the search for a particular solution of (3.45) as a conformal Lie derivative by the following:

$$\tilde{\mathcal{L}}_{\gamma,conf}X := \tilde{N}^{-1}\mathcal{L}_{\gamma,conf}X, \tag{3.70}$$

where $\tilde{N}$ is a given scalar, the unphysical densitized lapse defined by (3.14), which depends on the choice of γ.

We define an operator $\tilde{\Delta}_{\gamma,conf}$ by

$$(\tilde{\Delta}_{\gamma,conf}X)^j := D_i(\tilde{\mathcal{L}}_{\gamma,conf}X)^{ij}. \tag{3.71}$$

The mathematical properties of $\Delta_{\gamma,conf}$ and $\tilde{\Delta}_{\gamma,conf}$ are essentially the same.

We now decompose the tensor $\tilde{K}$ as follows:

$$\tilde{K}^{ij} \equiv (\tilde{\mathcal{L}}_{\gamma,conf}X)^{ij} + U^{ij}. \tag{3.72}$$

Under a change of conformal metric $\gamma = \theta^{\frac{4}{n-2}}\gamma'$ we have $\tilde{N} = \theta^{\frac{-2n}{n-2}}\tilde{N}'$, hence

$$(\tilde{\mathcal{L}}_{\gamma',conf}X)^{ij} = \theta^{-2(n+2)/(n-2)}(\tilde{\mathcal{L}}_{\gamma,conf}X)^{ij}, \tag{3.73}$$

which is the required scaling.

The unknowns in the conformally formulated constraints are still X and φ, but the CF momentum constraint now reads:

$$\mathcal{M}^i \equiv (\tilde{\Delta}_{\gamma,conf}X)^j + D_iU^{ij} - \left\{\frac{n-1}{n}\varphi^{2n/(n-2)}\gamma^{ij}\partial_j\tau + \tilde{J}_1^i + \tilde{J}_0^i\varphi^{4/(n-2)}\right\} = 0. \tag{3.74}$$

We have then the following corollary to Theorem 3.4 under otherwise the same hypothesis.

Corollary 3.7 *If the tensor $\tilde{K}$ solution of the CF momentum constraint in a metric γ is obtained as the sum of a given traceless tensor U and the product by a given function $\tilde{N}^{-1}$ of the conformal Lie derivative of a vector X:*

$$\tilde{K}^{ij} \equiv (\tilde{\mathcal{L}}_{\gamma,conf} X)^{ij} + U^{ij}, \tag{3.75}$$

then the tensor

$$\tilde{K}'^{ij} \equiv (\tilde{\mathcal{L}}_{\gamma',conf} X)^{ij} + U'^{ij}, \quad U'^{ij} = \theta^{-2(n+2)/(n-2)} U^{ij}, \quad \tilde{N}' = \theta^{\frac{2n}{n-2}} \tilde{N} \tag{3.76}$$

is solution of the momentum constraint, conformally formulated in the metric γ'.

3.6.3 *Physical interpretation*

The physical solution $(\bar{g}, K)$ deduced from the conformally formulated one is such that the traceless part K_T of K is the following one:

$$K_T^{ij} = \varphi^{-2(n+2)/(n-2)} \{\tilde{N}^{-1} (\mathcal{L}_{\gamma,conf} X)^{ij} + U^{ij}\} \tag{3.77}$$

We compare this expression with the original expression of K in terms of the shift β, lapse $\bar{N}$ and partial derivative $\partial_t g_{ij}$, which is (see VI.3.14, VI.3.11)

$$K^{ij} = \frac{1}{2} \bar{N}^{-1} (\overline{\partial_t g^{ij}} + (\bar{L}_\beta \bar{g})^{ij}), \quad \text{with} \quad (\bar{L}_\beta \bar{g})^{ij} \equiv \bar{\nabla}^i \beta^j + \bar{\nabla}^j \beta^i, \tag{3.78}$$

from which it results that:

$$\tau := \bar{g}_{ij} K^{ij} = -\frac{1}{2} \bar{N}^{-1} (\overline{\partial_t \log(\det \bar{g})} + 2 \bar{\nabla}_i \beta^i) \tag{3.79}$$

We deduce from these formulae, by elementary calculus:

$$K_T^{ij} = \frac{1}{2} \bar{N}^{-1} \{(\mathcal{L}_{\bar{g},conf} \beta)^{ij} + (\det \bar{g})^{-\frac{1}{n}} \overline{\partial_t (\bar{g}^{ij} (\det \bar{g})^{\frac{1}{n}})}\}. \tag{3.80}$$

We use the relations (3.14) and (3.69) between $\bar{N}$ and $\tilde{N}$, $\mathcal{L}_{\bar{g},conf}\beta$, and $\mathcal{L}_{\gamma,conf}\beta$ to see that if we compute the initial t-derivatives $\overline{\partial_t g^{ij}}$ for a spacetime with the initial data $(\bar{g}, K)$ and source data U constructed with the conformal metric γ and shift β such that

$$\beta = 2X \tag{3.81}$$

we find that these derivatives $\overline{\partial_t g^{ij}}$ are such that

$$\frac{1}{2} \bar{N}^{-1} (\det \bar{g})^{-\frac{1}{n}} \overline{\partial_t [g^{ij} (\det \bar{g})^{\frac{1}{n}}]} = \varphi^{-2(n+2)/(n-2)} U^{ij}; \tag{3.82}$$

that is,

$$\overline{\partial_t [g^{ij} (\det \bar{g})^{\frac{1}{n}}]} = 2 \tilde{N} (\det \gamma)^{\frac{1}{n}} U^{ij}. \tag{3.83}$$

These formulae are in agreement with the conformal scaling (3.68) of U. The particular choice $U = 0$ leads to the stationary initial value $\overline{\partial_t g^{ij}} = 0$, with

a shift of the time lines given by (3.81). We express this result as a theorem that is physically important, since the traceless tensor U^{ij} with some conformal weight, acts as a source of the Hamiltonian–Lichnerowicz equation and may be considered, therefore, as representing in some sense the gravitational radiation at initial time, radiation which is zero in the initially stationary case.

Theorem 3.8 *(York) If in the thin sandwich conformal decomposition of $\tilde{K}$ the initially given traceless 2-tensor U vanishes, then the spacetime is initially stationary with respect to time lines with shift solution of the Equations (3.74).*

4 Case $n = 2$

The vacuum Einstein equations are trivial in the case $n = 2$ in the sense that $Ricc(\mathrm{g}) = 0$ implies that the spacetime metric g is locally flat when $n + 1 = 3$. However, the $2 + 1$-dimensional Einstein theory has a topological content: (V, g) is not necessarily the Minkowski spacetime M_3. In particular $V = S \times R$ with S a two-dimensional compact surface can be a Lorentzian flat manifold with S the torus T^2 or a surface of genus greater than 1. On the other hand 2+1-dimensional, non-flat, Einstein equations with sources appear for spacetimes which admit a one-parameter spacelike isometry group.

The constraints can be conformally formulated in the case $n = 2$ by setting[18]

$$\bar{g} = exp(2\lambda)\gamma \quad \text{and} \quad K^{ij} = exp(4\lambda)\tilde{K}^{ij} + \frac{1}{2}\bar{g}^{ij}\tau.$$

Then

$$\bar{\nabla}_i K^{ij} \equiv exp(4\lambda) D_i \tilde{K}^{ij} + \frac{1}{2}\bar{g}^{ij}\partial_i \tau$$

and

$$\bar{\nabla}_i K^{ij} - \bar{g}^{ij}\partial_i \tau \equiv exp(4\lambda) D_i \tilde{K}^{ij} - \frac{1}{2}\bar{\gamma}^{ij} exp(-2\lambda)\partial_i \tau.$$

The constraints split again into a linear system for $\tilde{K}$ and a semi-linear equation for λ when $\tilde{K}$ is known, if τ is a given constant and the momentum of the sources is zero, or properly weighted.

5 Solutions on compact manifolds

In the following sections we will study the solution of the conformally formulated constraints on an initial Riemannian space n-manifold (M, γ), an elliptic system, as it should be for the determination of data at an instant of time. The problem to solve is a global problem on (M, γ). Its resolution depends crucially on the global properties of the manifold (M, γ). Of course the solution depends on the data for the sources, called $q's$ and $J's$ in our equations. The solution depends also on τ, which will be the mean curvature of the initial manifold in the ambient

[18] Studied in Moncrief, V. (1986) *Ann. Phys.*, **167**, 118–42.

spacetime, and on a symmetric traceless tensor U, which, roughly speaking[19], represents the initial radiation field, whose energy contributes to curvature. In the thin sandwich conformal formulation one gives in addition a scalar $\tilde{N}$. All these data must be chosen to obtain initial data modelling – as best as possible – a specific physical problem.

We will first study the case of a *compact manifold*[20] M (without boundary as always supposed when not specified otherwise), of arbitrary dimension, which is important for cosmological applications and perhaps quantization, and which is mathematically interesting.

When M is a compact manifold, we endow it with a smooth Riemannian metric e to define functional spaces, in particular Sobolev spaces, independently of the metric γ conformal to the physical metric, because we do not want γ to be necessarily smooth. We denote by W^p_s (H_s if $s = 2$) the usual Sobolev space on (M, e), by M^p_s (H_s if $s = 2$) the space of W^p_s Riemannian metrics, open cone in W^p_s if $s > \frac{n}{p}$ (see Appendix I). Elliptic systems on a smooth compact Riemannian manifold (M, e) are studied in Appendix II. We recall that, unless otherwise stated, the manifold M is supposed to be oriented.

We treat the case of low differentiability W^p_2, $p > \frac{n}{2}$, which allows for discontinuities of the sources, and the case H_s, $s > \frac{n}{2}$, which matches with the functional spaces used in evolution problems. Both spaces contain the space H_2 when $n \leq 3$. We write up the case $n \geq 3$. Most results extend to the case $n = 2$, with only a change in notation. This is not so when M is an asymptotically Euclidean manifold[21].

6 Solution of the momentum constraint

The momentum constraint reads as the following equation for the vector field X:

$$(\Delta_{\gamma,conf} X)^i \equiv D_j(\mathcal{L}X)^{ij} = F^i, \tag{6.1}$$

$$F^i \equiv \frac{n-1}{n}\varphi^{2n/(n-2)}\gamma^{ij}\partial_j\tau + \varphi^{4/(n-2)}\tilde{J}^i_0 + \tilde{J}^i_1 - D_j U^{ij} \tag{6.2}$$

When γ, a Riemannian metric, U, a given traceless symmetric 2-tensor, τ, $\tilde{J}_1$, $\tilde{J}_0$, and φ are known, (6.1) is a linear system for the vector X.

Lemma 6.1 *If $\gamma \in M^p_s$, $s > \frac{n}{p}$, $s \geq 2$, the vector F belongs to W^p_{s-2}, as soon as $\varphi \in W^p_s$, $D\tau, \tilde{J} \in W^p_{s-2}$ and $U \in W^p_{s+1}$.*

[19] Rigorously speaking, "radiation" in General Relativity cannot be separated from other energy sources.

[20] See Choquet-Bruhat, Y. (2004) *Class. Quant. Grav.*, **21**(3), 127–51, and references therein.

[21] The Laplacian of Euclidean space has very different properties in two and higher dimension. Its elementary solution is r^{-1} for $n = 3$; it is positive and tends to zero at infinity. None of these properties holds for $\log r$, the elementary solution for $n = 2$.

Proof The Sobolev multiplication theorem (see Appendix I). □

Theorem 6.2 *The momentum constraint (6.1) with $\gamma \in M_2^p$, $p > \frac{n}{2}$, has a solution $X \in W_2^p$, if $F \in L^p$ and if F is $L^2(\gamma)$ orthogonal to the space of conformal Killing (CK) vector fields of (M, γ).*

The solution is determined up to addition of a CK vector. The solution is unique if we impose that it be $L^2(\gamma)$ orthogonal to CK vectors; in this case, there exists a constant $C_\gamma > 0$ depending only on γ such that

$$||X||_{W_2^p} \leq C_\gamma ||F||_{L^p}. \tag{6.3}$$

The theorem holds with M_2^p replaced by M_s, W_2^p by H_s, and L^p replaced by H_{s-2}, $s > \frac{n}{2}$.

Proof This theorem results from the previous lemma and from Theorem x.2.3 in Appendix II when (M, γ) admits no conformal Killing vector fields.

The extension to the case of the existence of such CK vectors is an exercise in functional analysis which may proceed as follows. The equality (2.44) in Appendix II gives that

$$\int_M X_j (\Delta_{\gamma,conf} X - kX)^j \mu_\gamma = -\int_M \{|\mathcal{L}_{\gamma,conf} X|_\gamma^2 + k|X|_\gamma^2\} \mu_\gamma \tag{6.4}$$

Therefore $\Delta_{\gamma,conf} - k : W_2^p \to L^p$ is injective, and hence an isomorphism, if k is a strictly positive number. Its inverse is then a compact operator (see Appendix 1); hence $\Delta_{\gamma,conf}$ is a Fredholm operator, and $\Delta_{\gamma,conf} X = F$ has a solution $X \in W_2^p$ iff F is orthogonal to the kernel of the adjoint operator, which is $\Delta_{\gamma,conf}$ itself. The solution is unique in the Banach space of W_2^p vectors orthogonal to CK vectors.

The proof of higher regularity for more regular coefficients, and of the inequality (6.3), is obtained by differentiating the equations. □

Corollary 6.3 *The same theorem holds if $\Delta_{\gamma,conf}$ is replaced by $\tilde{\Delta}_{\gamma,conf} := D_i \tilde{\mathcal{L}}_{\gamma,conf}$, with $\tilde{\mathcal{L}}_{\gamma,conf} := \tilde{N}^{-1} \mathcal{L}_{\gamma,conf}$, where $\tilde{N} \in W_{2+s}^p$ a given function such that $\tilde{N} > 0$.*

Proof The operators $\Delta_{\gamma,conf}$ and $\tilde{\Delta}_{\gamma,conf}$ have the same principal part, except for the product by the positive and continuous (under the hypothesis $s \geq 0$, $p > \frac{n}{2}$) function $\tilde{N}^{-1}$. They have the same kernel, the space of conformal Killing vector fields, because integration by parts gives, as in the appendix,

$$\int_M X_j (\tilde{\Delta}_{\gamma,conf} X)^j \mu_\gamma \equiv \int_M X_j D_i \left\{ \tilde{N}^{-1} \left(D^i X^j + D^j X^i - \frac{2}{n} \gamma^{ij} D_k X^k \right) \right\} \mu_\gamma$$

$$= -\int_M \tilde{N}^{-1}\left(D^iX^j + D^jX^i - \frac{2}{n}\gamma^{ij}D_kX^k\right)$$
$$\times\left(D_iX_j + D_jX_i - \frac{2}{n}\gamma_{ij}D_lX^l\right)\mu_\gamma. \tag{6.5}$$

An analogous proof shows that $\tilde{\Delta}_{\gamma,conf}$, like $\Delta_{\gamma,conf}$, is self-adjoint. □

Remark 6.4 *When X is not unique, it is determined up to the addition of a CK vector field. The tensor $\tilde{K}$, given by the formula (3.72), is still unique.*

Remark 6.5 *It is known from ellipticity that on a compact manifold the kernel of $\Delta_{\gamma,conf}$ is finite-dimensional; it is in fact of dimension at most $\frac{1}{2}(n+1)(n+2)$ on an $n > 1$-dimensional manifold.*

Remark 6.6 *If the Ricci tensor of γ, with components ρ_{ij}, is negative definite the manifold (M,γ) admits no* ***conformal Killing fields****. Indeed the equality $\Delta_{\gamma,conf}X = 0$ implies, using the Ricci identity,*

$$D^iD_iX_j + \left(1 - \frac{2}{n}\right)D_jD^iX_i + \rho_{jl}X^l = 0, \tag{6.6}$$

which implies on a compact manifold

$$\int_M \left\{-D^iX^jD_iX_j - \left(1 - \frac{2}{n}\right)D_jX^jD^iX_i + \rho_{jl}X^lX^j\right\}\mu_\gamma = 0, \tag{6.7}$$

from which follows $X \equiv 0$ if Ricci(γ) is negative definite and $n \geq 2$.

The generic absence of CK fields has been established by Chruściel, Beig, and Schoen[22].

Lemma 6.7 *When $D\tau \equiv 0$ and $J_0 \equiv 0$, then F is L^2 orthogonal to the space of CK vector fields if and only if it is so of J_1.*

Proof If $D\tau \equiv 0$ and $J_0 \equiv 0$, then F reduces to the sum of J_1 and the divergence of a traceless symmetric tensor. Such divergences are L^2 orthogonal to CK vector fields. This fact, as is well known in smoother cases, results from the identity obtained by integration by parts and is still valid under our hypotheses, when the tensor U is traceless,

$$\int_M D_iU^{ij}X_j\mu_\gamma = -\int_M U^{ij}(\mathcal{L}_{\gamma,conf}X)_{ij}\mu_\gamma. \tag{6.8}$$

□

If $D\tau \not\equiv 0$ or $J_0 \not\equiv 0$ (matter sources with non-zero momentum or scalar field with non-zero potential) the orthogonality condition of F to CK vector fields depends on φ, raising a new difficulty for the coupling with the Lichnerowicz equation.

[22] Chruściel, P., Beig, R., and Schoen, R. *Adv. Theor. Math. Phys.*, gr-qc/0403042.

7 Lichnerowicz equation

The Lichnerowicz equations are of the type

$$\mathcal{H} \equiv \Delta_\gamma \varphi - f(.,\varphi) = 0, \qquad f(.,\varphi) \equiv r\varphi - \sum_I a_I \varphi^I = 0, \tag{7.1}$$

The exponents I are some positive or negative real numbers depending on the nature of the sources and their scalings. The coefficients r, a_I depend on the mean curvature τ, on the solution $\tilde{K}$ of the momentum constraint, on the sources, and on their scalings.

When a scalar field $\bar{\psi}$ is present r is given by

$$r \equiv k_n[R(\gamma) - |D\bar{\psi}|^2_\gamma], \quad k_n = \frac{n-2}{4(n-1)}. \tag{7.2}$$

For the Einstein-scalar field equations in vacuum with potential V, the additional term reduces to

$$\sum_I a_I \varphi^I \equiv k_n \left\{ |\tilde{K}|^2_\gamma + |\tilde{\pi}|^2) \varphi^{-\frac{3n-2}{n-2}} - \left[\frac{n-2}{4n}\tau^2 - V(\bar{\psi}) \right] \varphi^{\frac{n+2}{n-2}} \right\}. \tag{7.3}$$

7.1 *The Yamabe classification*

7.1.1 *The Yamabe theorem*

A theorem conjectured by Yamabe and proved by Trudinger, Aubin, and Schoen in an increasing (and ultimately exhaustive) number of cases[23] has been used by Isenberg to classify solutions obtained by the conformal method in the case of smooth given metrics γ. We will use here a weaker version, Theorem 7.6, which applies to less regular metrics. We first prove the extension to metrics γ in M^p_2, $p > \frac{n}{2}$, of some lemmas, well known for smooth metrics .

Lemma 7.1 *If $\gamma \in M^p_2$, $p > \frac{n}{2}$, the functional, called a q-Yamabe functional,*

$$J_{\gamma,q}(\varphi) \equiv \frac{\int_M (k_n^{-1}|\nabla\varphi|^2_\gamma + R(\gamma)\varphi^2)\mu_\gamma}{(\int_M \varphi^{2q}\mu_\gamma)^{1/q}}. \tag{7.4}$$

is defined for every $\varphi \in H_1$, $\varphi \not\equiv 0$, $1 \le q \le n/(n-2)$.

Proof If $\gamma \in M^p_2$, $p > \frac{n}{2}$, then $\gamma \in C^0$, and is uniformly equivalent to e. Therefore the pointwise norms of a tensor in the metrics γ and e are uniformly equivalent, and their L^p norms in the metrics γ and e are equivalent. The Sobolev embedding and multiplication theorems (see Appendix I) show that $R(\gamma) \in L^p$. On the other hand, if $\varphi \in H_1$, then $\varphi \in L^{2q}$, $1 \le q \le n/(n-2)$, hence $R(\gamma)\varphi^2 \in L^1$ if $n \ge 2$. □

Note that the embedding of H_1 into L^{2q} is compact for $q < n/(n-2)$ but not for $q = n/(n-2)$. This non compactness is the origin of the incorrectness of Yamabe's proof.

[23] See for instance CB-DM2 supplement 6 and references therein.

Lemma 7.2 *The $n/(n-2)$-Yamabe functional, called simply the Yamabe functional and denoted $J_\gamma(\varphi)$, is a conformal invariant in the following sense*

$$J_{\gamma'}(\varphi') = J_\gamma(\varphi), \quad \text{if} \quad \gamma' = \theta^{\frac{4}{n-2}}\gamma, \quad \varphi' = \theta^{-1}\varphi. \tag{7.5}$$

Proof The relations between γ and γ', φ and φ' imply, since $\mu_\gamma = \theta^{-\frac{2n}{n-2}}\mu_{\gamma'}$,

$$\int_M \varphi^{2n/(n-2)}\mu_\gamma = \int_M \varphi'^{2n/(n-2)}\mu_{\gamma'}, \tag{7.6}$$

while integration by parts of (3.55) implies by straightforward calculation

$$\int_M (k_n^{-1}|\nabla\varphi|^2_\gamma + R(\gamma)\varphi^2)\mu_\gamma = \int_M (k_n^{-1}|\nabla\varphi|^2_{\gamma'} + R(\gamma')\varphi^2)\mu_{\gamma'} \tag{7.7}$$

□

Lemma 7.3 *Given $\gamma \in M_2^p$, $p > \frac{n}{2}$, a functional $J_{\gamma,q}(\varphi)$ admits an infimum*[24] *for $\varphi \in W_2^p$ and $\varphi \not\equiv 0$. In the case where $q = n/(n-2)$, this infimum depends only on the conformal class of γ. It is called the Yamabe invariant.*

Proof If $\gamma \in M_2^p$, $p > \frac{n}{2}$, then, on the compact manifold M, $R(\gamma) \in L^r$ for all $1 \leq r \leq p$, hence $||R(\gamma)||_{L^{q/(q-1)}}$ is bounded if $q \leq n(n-2)$. The following inequality, which is a consequence of the Hölder inequality,

$$\int_M (k_n^{-1}|\nabla\varphi|^2 + R(\gamma)\varphi^2)\mu_\gamma \geq -\left|\int_M R(\gamma)\varphi^2\mu_\gamma\right| \geq -||R(\gamma)||_{L^{q/(q-1)}}||\varphi^2||_{L^q}. \tag{7.8}$$

shows then that the functional $J_{\gamma,q}$ is bounded below, and therefore admits an infimum μ_q. It can be proved (see Aubin, 1982, p. 128) that μ_q is a continuous function of q, everywhere negative, everywhere zero, or everywhere positive for $1 \leq q \leq n/(n-2)$. □

The infimum $\mu_{n/(n-2)}$, denoted simply μ, is called the **Yamabe number**.

$$\mu \equiv \underset{\varphi\in W_2^p, \varphi\not\equiv 0}{Inf} J_\gamma(\varphi). \tag{7.9}$$

Lemma 7.4 *The Yamabe number depends only on the conformal class of γ.*

Proof The lemma 7.2. □

The manifolds (M, γ) are split into three **Yamabe classes** according to the sign of μ.

Definition 7.5 *The manifold (M, γ) is said to be in the negative Yamabe class if $\mu < 0$, in the zero Yamabe class if $\mu = 0$, in the positive Yamabe class if $\mu > 0$.*

[24] Recall that an infimum is the largest number μ such $J(\varphi) \geq \mu$ for all φ in the considered class, it is not necessarily attained by some function φ_m.

It results from the definition that a sufficient condition for (M, γ) to be of negative Yamabe class is $\int_M R(\gamma)\mu_\gamma < 0$, since then $J_\gamma(1) < 0$.

The following theorem, which is an extension to M_2^p metrics of theorems proved by Trudinger and Aubin[25] for smooth metrics, will be sufficient for our use.

Theorem 7.6 *Let (M, γ) be a compact Riemannian manifold with $\gamma \in M_2^p$, $p > \frac{n}{2}$. Then:*

1. *If γ is in the negative Yamabe class it is conformal to a metric with scalar curvature -1.*
2. *If γ is in the zero Yamabe class it is conformal to a metric with scalar curvature 0.*
3. *If γ is in the positive Yamabe class, it is conformal to a metric with continuous and positive scalar curvature.*

Proof Yamabe proved, by variational methods (construction of a minimizing sequence) and functional analysis (using in particular the compactness of the embedding H_1 into L^{2q}, $q < n/(n-2)$), that the infimum μ_q of a q-Yamabe functional is attained by a function $\varphi_{m,q} \in H_1$, $\varphi_{m,q} \not\equiv 0$, if $1 \leq q < n/(n-2)$. The function $\varphi_{m,q}$ is a critical point[26] of the functional, hence it satisfies in a weak sense the Euler–Lagrange equation of the q-Yamabe functional, which we refer to as the q-**Yamabe equation**

$$k_n^{-1}\Delta_\gamma \varphi_{m,q} - R(\gamma)\varphi_{m,q} + \mu_q \varphi_{m,q}^{2q-1} = 0. \tag{7.10}$$

One can show, using elliptic theory, that $\varphi_{m,q} \in W_2^p$ and $\varphi_{m,q} > 0$.

The case $q = n/(n-2)$ ("critical exponent") is significantly more delicate, in particular with respect to showing $\varphi_m := \varphi_{m,q} > 0$. The embedding of H_1 into $L^{2n/n-2}$ is not compact. We sketch the proof in this case.

1. Case $\mu < 0$. Then $\mu_q < 0$ for $1 \leq q < n/(n-2)$. It is possible to show[27] that the functions $\varphi_{m,q}$, $q < n/(n-2)$ are uniformly bounded in W_2^p, converge in H_1 to a function φ_m which satisfies the Yamabe equation, is in W_2^p, and strictly positive. Then $\gamma_m \equiv \varphi_m^{4/(n-2)}\gamma$ is a Riemannian metric. It has constant scalar curvature μ given by (see the identity 3.1) the equality

$$R(\gamma_m) = \mu. \tag{7.11}$$

By rescaling of γ_m one obtains a metric γ' with scalar curvature -1 since the scalar curvature of cg, c a positive number, is $c^{-1}g$.

[25] Trudinger, N. (1968) *Ann. Norm. Sup. Pisa*, **22**, 265–74; Aubin, T. J. (1976) *Math. Pures Appl.*, **55**, 269–96.

[26] That is, the Frechet derivative of the functional vanishes at that point.

[27] See for instance CB-DM2, supplement 6, p. 491.

2. If $\mu = 0$, then also $\mu_q = 0$. Each metric $\gamma_{m,q} \equiv \varphi_{m,q}^{4/(n-2)}\gamma$ satisfies the equation

$$k_n^{-1}\Delta_\gamma \varphi_{m,q} - R(\gamma)\varphi_{m,q} = 0, \tag{7.12}$$

and hence has scalar curvature zero.

3. If $\mu > 0$, hence also $\mu_q > 0$, we write the equation satisfied by the metric $\gamma_{m,q} \equiv \varphi_{m,q}^{4/(n-2)}\gamma$ under the form

$$k_n^{-1}\Delta_\gamma \varphi_{m,q} - R(\gamma)\varphi_{m,q} + \mu_q \varphi_{m,q}^{(2q-1)-(n+2)/(n-2)}\varphi_{m,q}^{(n+2)/(n-2)} = 0, \tag{7.13}$$

and we see that the scalar curvature of $\gamma_{m,q}$ is the continuous positive function $\mu_q \varphi_{m,q}^{(2q-1)-(n+2)/(n-2)}$. □

The following theorem is much more difficult to prove.

Theorem 7.7 *If γ is a smooth metric in the positive Yamabe class, then γ is conformal to a metric with scalar curvature $+1$.*

The proof was given by Aubin[28] in the case $n \geq 6$ and γ not locally conformally flat. It involves estimates of the "best Sobolev constant", using geodesic balls. The proofs for γ possibly conformally flat, and for the case $6 > n \geq 3$ are due to Schoen[29]. It involves in the last case the estimate of the second term in the expansion of the Green function; this estimate is related to the positive mass theorem of General Relativity. These proofs, for which we refer the reader to the original articles, require at least a $C^{1,1}$ metric, implied only by $\gamma \in M_{2+s}^p$, $p > \frac{n}{2}$ and $s \geq 2$. The existence of a constant scalar curvature metric in the positive Yamabe class is not proved for $\gamma \in M_2^p$ metrics.

7.1.2 *Topology and Yamabe classification*

An interesting question is: which Yamabe class of metrics can a given manifold support? The following theorem shows that the existence on M of a metric in the negative Yamabe class does not restrict the topology of M.

Theorem 7.8 *(Aubin, 1970)*[30] *Every compact manifold of dimension $n \geq 3$ admits a metric in the negative Yamabe class.*

Proof See Aubin (1982). □

On the other hand, not all manifolds can support metrics in the positive[31] or zero Yamabe classes.

[28] See references in Aubin, T. (1982) *Non-Linear Analysis and Monge-Ampère Equations*, Springer, London.

[29] Schoen, R. (1984) *J. Diff. Geom.*, **20**, 479–95.

[30] Aubin, T. (1970) *J. Diff. Geom.*, **4**, 383–424; see also Aubin, T. (1982) *Non-Linear Analysis on Manifolds and Monge–Ampère Equations*, Springer, and references therein.

[31] The first example of manifolds which do not admit a metric in the positive Yamabe class was given by Lichnerowicz, A. (1963) *C. R. Acad. Sci. Paris*, **7**.

To link the Yamabe properties with the topology of M, Fisher and Moncrief[32] introduce the following **Yamabe types**, which are mutually exclusive, because[33] the absence of a metric in the zero Yamabe class implies the non-existence of a metric in the positive Yamabe class. The proof of this last fact uses the property $R(g) = c^{-1}R(cg)$ and a continuity argument.

Definition 7.9 *Let M be a compact manifold of dimension $n \geq 3$. It is said to be:*

1. *Of positive Yamabe type if it admits a metric of positive Yamabe class.*
2. *Of zero Yamabe type if it admits a metric in the zero Yamabe class, but no metric in the positive Yamabe class.*
3. *Of negative Yamabe type if it admits no metric in the zero or positive Yamabe class*[34].

It results from this definition and the expression of the Yamabe number that:

1. If M is of negative Yamabe type, then all metrics on M are in the negative Yamabe class, hence every scalar curvature function (i.e. curvature of some metric on M) is negative somewhere.

2. If M is of zero Yamabe type, then all metrics on M are in the negative or zero Yamabe class, hence any curvature function is either negative somewhere or is identically zero.

We have already mentioned that the topology of manifolds of positive or zero Yamabe type is not arbitrary.

For any dimension $n \geq 3$, Gromov and Lawson[35] have proven that connected sums of spherical space forms (manifolds quotients of S^n by the action of a finite isometry group) and products $S^{n-1} \times S^1$ are of the positive Yamabe type. The torus T^n is a manifold of zero Yamabe type.

In the case $n = 3$ Gromov and Lawson have obtained a converse result, using in particular the Poincaré conjecture, now proved by Perelman following work by Hamilton, that all smooth, orientable, simply connected 3-manifolds are diffeomorphic to the sphere S^3. If not simply connected all smooth, orientable 3-manifolds of the positive Yamabe type are connected sums of spherical space forms, manifolds finitely covered by S^3, and products $S^2 \times S^1$.

[32] Fisher, A. and Moncrief, V. (1994) in *Physics on Manifolds* (eds. M. Flato, R. Kerner and A. Lichnerowicz), Kluwer, pp. 111–51.

[33] Kazdan, J. (1985) Prescribing the curvature of a Riemannian manifold. *Regional Conference Series in Maths*, No. 57, AMS, Providence RI, and references therein.

[34] In fact the absence of a metric in the zero Yamabe class implies the non-existence of metrics in the positive Yamabe class (see Aubin and Kazdan, who use the fact that $R(cg) = c^{-1}$ $R(cg)$ and a continuity argument).

[35] Gromov, M. and Lawson, H. B. (1980) *Ann. Math.*, **111**, 423–34.

It has been proved by Aubin[36] that the Yamabe numbers of any conformal class on a manifold M are bounded above by the Yamabe number of the conformal class of the standard metric on the sphere S^3. The Yamabe numbers of conformal classes of metrics on a given manifold M admit therefore a supremum, usually denoted $\sigma(M)$ and called the **σ-constant**. This constant is a topological invariant of the manifold M.

The σ-constant has been used by Fisher and Moncrief[37] in their construction of reduced Hamiltonians on spacetimes with space manifolds of negative Yamabe type.

7.1.3 *Scalar extended Yamabe theorem*

We show briefly how the Yamabe classification can be extended in the presence of a scalar field[38].

When a scalar field ψ is present among the sources, the coefficient of the linear term in φ in the Lichnerowicz equation is

$$r = k_n[R(\gamma) - |D\bar{\psi}|^2_\gamma]. \tag{7.14}$$

We suppose that $\gamma \in M_2^p$, $D\bar{\psi} \in W_1^p$, $p > \frac{n}{2}$, and we consider the functional, defined for $\varphi \in W_2^p$, $\varphi \not\equiv 0$:

$$J_{\gamma,\bar{\psi}}(\varphi) \equiv \frac{\int_M \{k_n^{-1}|\nabla\varphi|^2_\gamma + [R(\gamma) - |D\bar{\psi}|^2_\gamma]\varphi^2)\mu_\gamma}{(\int_M \varphi^{2n/(n-2)}\mu_\gamma)^{(n-2)/n}}. \tag{7.15}$$

Lemma 7.10

1. *$J_{\gamma,\bar{\psi}}$ is bounded below.*
2. *Its infimum is a conformal invariant.*

Proof

1. The proof of the lemma 7.1 holds, with $R(\gamma)$ replaced by $R(\gamma) - |D\bar{\psi}|^2_\gamma$ because $W_1^p \subset L^{2p}$ when $p > \frac{n}{2}$, hence $D\bar{\psi} \in W_1^p$ implies $|D\bar{\psi}|^2_\gamma \in L^p$.

2. Under a change of conformal metric

$$\gamma' = \theta^{\frac{4}{n-2}}\gamma \tag{7.16}$$

it holds that

$$|D\bar{\psi}|^2_\gamma \equiv \gamma^{ij}\partial_i\bar{\psi}\partial_j\bar{\psi} = \theta^{\frac{4}{n-2}}\gamma'^{ij}\partial_i\bar{\psi}\partial_j\bar{\psi} \equiv \theta^{\frac{4}{n-2}}|D\bar{\psi}|^2_{\gamma'}. \tag{7.17}$$

Hence

$$J_{\gamma,\bar{\psi}}(\varphi) = J_{\gamma',\bar{\psi}}(\varphi') \quad \text{with} \quad \varphi' = \theta^{-1}\varphi. \tag{7.18}$$

[36] See Aubin (1982).

[37] Fisher, A. and Moncrief, V. (1998) in *Mathematical and Quantum aspects of General Relativity and Cosmology* (eds. S. Cotsakis and G. Gibbons), Lecture Notes in Physics 537, Springer, London, pp. 70–102.

[38] Choquet-Bruhat, Y., Isenberg, J., and Pollack, D. (2007) *Class. Quant. Grav.*, **24**(1), 800–29.

We say that a pair $(\gamma, \bar{\psi})$ is in the positive, negative, or zero scalar extended Yamabe class, abbreviated to SY class, if this invariant is positive, negative, or zero. The arguments used to prove Lemma 7.3 show that it extends to the new definition of r, for a given scalar $\bar{\psi}$. □

7.2 *Non-existence and uniqueness*

Let $f \geq 0$ be a function defined almost everywhere on M. We say that[39] $f \not\equiv 0$ if there is an open set of M where $Inf\ f > 0$.

Theorem 7.11 *(Non-existence) The Lichnerowicz equation (7.1) admits no solution $\varphi > 0$, $\varphi \in W_2^p$, $p > \frac{n}{2}$, on a compact manifold (M, γ), $\gamma \in M_2^p$, $a_I \in L^1$ if $r - \sum a_I \not\equiv 0$ and either:*

1. *$r \leq 0$, $a_I \geq 0$ for all I.*
2. *$r \geq 0$, $a_I \leq 0$ for all I.*

Proof We approximate the given γ and φ in $W_2^p \subset C^0$ by sequences $\gamma_n \in C^1$ and $\varphi_n \in C^2$. The integral of $\Delta_{\gamma_n}\varphi_n$ with respect to the volume element of γ_n is equal to zero on a compact manifold M, by the Stokes formula. Then the Sobolev embedding and multiplication theorems show that $\Delta_\gamma \varphi(\mu_e)^{-1}\mu_\gamma - \Delta_{\gamma_n}\varphi(\mu_e)^{-1}\mu_{\gamma_n}$ tends to zero in L^1 (see analogous proofs in Appendix II; remark that the volume element of γ_n tends in C^0 to the volume element of γ) hence the integral of $\Delta_\gamma \varphi$ on (M, γ) is zero.

On the other hand, $f(., \varphi)$ is integrable on M since $\varphi \in C^0$ and $\varphi > 0$. Its integral cannot vanish under condition 1 [respectively 2] which implies that $f(., \varphi) \leq 0$ [respectively ≥ 0] and is strictly negative [positive] on an open set of M. □

The geometrical origin of the Lichnerowicz equation leads to a general uniqueness theorem, independent of the sign of r, hence of the scalar extended Yamabe class of γ. Since the sign of the conformal weight associated to q_2, i.e. to the magnetic energy of the 2-form F, depends on the dimension n, we state as follows the uniqueness theorem, leaving its specific applications to the reader (applications are given in the sections on the asymptotically Euclidean case).

Theorem 7.12 *(Uniqueness) The equation*

$$\Delta_\gamma \varphi - r\varphi = -a_I \varphi^{P_I}, \quad \text{sum over } I, \quad |P_I| \geq 1 \tag{7.19}$$

on (M, γ), with $r \in L^p$ given by (7.14). $\gamma \in M_2^p, p > \frac{n}{2}$, $a_I \in L^1$, has at most one non-constant positive solution $\varphi \in W_2^p$, if $a_I \geq 0$ for $P_I \leq 0$ and $a_I \leq 0$ for $P_I \geq 0$.

[39] If f is defined everywhere, the definition coincides with the usual notation.

Proof Suppose it admits two solutions $\varphi_1 > 0$ and $\varphi_2 > 0$. Using the identity (3.56) we obtain, with $q = \frac{2}{n-2}$, that:

$$\Delta_{\varphi_2^{2q}\gamma}(\varphi_1\varphi_2^{-1}) - r(\varphi_2^{2q}\gamma)(\varphi_1\varphi_2^{-1}) \equiv -(\varphi_1\varphi_2^{-1})^{(n+2)/(n-2)} r(\varphi_1^{2q}\gamma). \tag{7.20}$$

Since φ_1 is a solution of (7.1) we have

$$r(\varphi_1^{2q}\gamma) \equiv -\varphi_1^{-(n+2)/(n-2)}\{\Delta_\gamma\varphi_1 - \varphi_1 r(\gamma)\} = \varphi_1^{-(n+2)/(n-2)} a_I \varphi_1^{P_I}$$

and an analogous equation for $r(\varphi_2^{2q}\gamma)$. Inserting in the previous equation gives an equation of the form

$$\Delta_{\varphi_2^{2q}\gamma}(\varphi_1\varphi_2^{-1} - 1) - \lambda\{(\varphi_1\varphi_2^{-1} - 1\} = 0 \tag{7.21}$$

with

$$\lambda \equiv a_I \varphi_1^{P_I} \varphi_2^{-(n+2)/(n-2)} \mu, \quad \mu := \frac{(\varphi_1\varphi_2^{-1})^{-P_I} - 1}{\varphi_1\varphi_2^{-1} - 1}$$

Since φ_1 and φ_2 are continuous and positive on M, it holds, by the mean value formula, that μ is a continuous function on M, positive if $P_I < 0$, negative $P_I > 0$. Therefore, under the given hypothesis, λ is an integrable function on M and $\lambda \geq 0$, with $\lambda \not\equiv 0$ on M if $\Sigma|a_I| \not\equiv 0$. The equation implies then[40] $\varphi_1\varphi_2^{-1} - 1 \equiv 0$. □

7.3 *Existence theorems*

Existence results can be obtained by using the Leray–Schauder degree, as was done in the case $n = 3$, in Choquet-Bruhat (1972)[41] (with unscaled sources) and O'Murchada and York (1973)[42] (with scaled sources). It is also possible to use a constructive method (Isenberg, 1995)[43]. Both methods use sub- and supersolutions. Some of the following theorems are extensions of results obtained before with higher regularity (for $n = 3$ see reviews in Choquet-Bruhat and York, 1980; Isenberg, 1995). The theorems include recent results of Choquet-Bruhat (2004)[44], Maxwell (2005)[45], and Choquet-Bruhat *et al.* (2006)[46].

In order not to make the analysis too heavy, we suppose that in the case $n > 3$ there are no magnetic sources. We set

$$q = q_1 + q_2 \quad \text{with} \quad q_2 \equiv 0 \quad \text{if} \quad n > 3. \tag{7.22}$$

The Lichnerowicz equation reduces then to

$$\Delta_\gamma\varphi - r\varphi + a\varphi^{\frac{-3n+2}{n-2}} + q\varphi^{\frac{-n}{n-2}} - b\varphi^{\frac{n+2}{n-2}}, \quad a, q \geq 0. \tag{7.23}$$

[40] See Appendix II, Lemma 2.6.
[41] Choquet-Bruhat, Y. (1972) *C. R. Acad. Sci. Paris*, **274**, 682–4.
[42] O'Murchada, N. and York, J. (1973) *J. Math.*, **14**, 1551–7.
[43] Isenberg, J. (1995) *Class. Quant. Grav.*, **12**, 22–49.
[44] Choquet-Bruhat, Y. (2004) *Class. Quant. Grav.*, **21**(3), S127–S151
[45] Maxwell, D. (2005) *J. Hyp. Diff. Eq.*, **2**, 521–46
[46] Choquet-Bruhat, Y., Isenberg, J., and Pollack, D. (2006) *Chinese Ann. Math.*, **27**, serie B, No. 1, 31–52

We will always suppose[47] that $\gamma \in M_2^p$, $p > \frac{n}{2}$. The coefficients a, q, b are assumed to be functions[48] in L^p. We also suppose most of the time that r is conformally transformed by the SY (Scalar extended Yamabe) theorem to a continuous function of constant sign (or zero).

We recall the existence theorem proved in Appendix II.

Theorem 7.13 *Consider the semilinear elliptic equation*

$$\Delta_\gamma \varphi = f(x, \varphi), \tag{7.24}$$

with $\gamma \in M_2^p$, and $f : M \times I \to R$ by $(x, y) \mapsto f(x, y)$ a function which is L^p in $x \in M$ and smooth in $y \in I \equiv [\ell, m]$, $m \leq +\infty$. Suppose that this equation admits a subsolution φ_- and a supersolution φ_+, $\ell \leq \varphi_- \leq \varphi_+$, functions in W_2^p such that almost everywhere on M

$$\Delta_\gamma \varphi_- \geq f(., \varphi_-), \qquad \Delta_\gamma \varphi_+ \leq f(x, \varphi_+). \tag{7.25}$$

Then Equation (7.24) admits on M a solution $\varphi \in W_2^p$, $\ell \leq \varphi_- \leq \varphi \leq \varphi_+$. This solution can be constructed by iteration.

If, in addition, $\gamma \in M_s$, and f is H_{s-2} in x with $s > \frac{n}{2}$, then $\varphi \in H_s$.

The results for the Lichnerowicz equation depend on the signs of the coefficients b and r (recall that always $a \geq 0$ and $q \geq 0$). Some results can be obtained by looking for constant sub- and supersolutions. However, such solutions do not always exist and moreover their existence often requires L^∞ bounds of the coefficients, while we like to require of them as little as we can, which in our case is that they belong to L^p. This can often be achieved by taking advantage of the conformal invariance of the Lichnerowicz equation, looking for sub- and supersolutions of a conformally transformed equation[49].

7.3.1 *Case $b \geq 0$*

This case always happens if there are no unscaled matter sources and the potential of the scalar field is non-positive.

Theorem 7.14 *Suppose that $\gamma \in M_2^p$, $p > \frac{n}{2}$, and that $r, b, a, q \in L^p$, $b \geq 0$. The Lichnerowicz Equation (7.1) admits a solution $\varphi > 0$, $\varphi \in W_2^p$ in the following cases.*

1. *$(M, \gamma, \bar{\psi})$ is in the positive SY class and $a + q \not\equiv 0$.*
2. *$(M, \gamma, \bar{\psi})$ is in the zero SY class, $b \not\equiv 0$ and $a + q \not\equiv 0$.*
3. *$(M, \gamma, \bar{\psi})$ is in the negative SY class, and $\underset{M}{Inf} b > 0$.*

If, in addition, $\gamma \in M_s$, $r, b, a, q \in H_{s-2}$, with $s > \frac{n}{2}$, then $\varphi \in H_s$.

[47] D. Maxwell takes $\gamma \in M_s^2$, $s > \frac{n}{2}$. Both $W_2^p, p > \frac{3}{2}$, and H_s, $s > \frac{3}{2}$ and integer, contain H_2 when $n = 3$. Both have their advantages when $n > 3$.

[48] When $n = 3$, the assumption $\gamma \in M^2_{\frac{3}{2}+\varepsilon}$, $K \in H_{\frac{1}{2}+\varepsilon}$ implies $K.K \in H_{-\frac{1}{2}+\varepsilon}$, a generalized function (also called a distribution).

[49] Such a procedure was already indicated in Choquet-Bruhat and York (1980), and developed in Isenberg, J. (1987) *Class. Quant. Grav.*, **12**, 12–49.

Proof We use an idea of D. Maxwell to conformally transform the Lichnerowicz equation into an equation which admits constant sub- and supersolutions.

1. By the change of conformal metric

$$\gamma' = \theta^{\frac{4}{n-2}}\gamma \tag{7.26}$$

we have

$$r(\gamma') = \theta^{-\frac{n+2}{n-2}}(-\Delta_\gamma\theta + r(\gamma)\theta). \tag{7.27}$$

We choose θ such that, with $k \geq 0$, $k \not\equiv 0$ to be chosen later,

$$\Delta_\gamma\theta - k\theta = -(a+q). \tag{7.28}$$

Since $a+q \geq 0$, and $a+q \not\equiv 0$ there exists such a $\theta > 0$, $\theta \in W_2^p$. With such a θ it holds that

$$r(\gamma') = \theta^{-\frac{n+2}{n-2}}[-k\theta + r(\gamma)\theta + a + q]. \tag{7.29}$$

The conformally transformed Lichnerowicz equation is, with $\varphi' = \theta^{-1}\varphi$,

$$\theta^{\frac{n+2}{n-2}}\Delta_{\gamma'}\varphi' - [-k\theta + r(\gamma)\theta + a + q]\varphi' + a'\varphi'^{\frac{-3n+2}{n-2}} + q'\varphi'^{\frac{-n}{n-2}} - b'\varphi'^{\frac{n+2}{n-2}} = 0, \tag{7.30}$$

with

$$a' = a\theta^{\frac{-3n+2}{n-2}}, \quad q' = q\theta^{\frac{-n}{n-2}}, \quad b' = b\theta^{\frac{n+2}{n-2}}. \tag{7.31}$$

The number Y is a constant supersolution of Equation (7.30) if:

$$[k\theta - r(\gamma)\theta - a - q]Y + a'Y^{\frac{-3n}{n-2}} + q'Y^{\frac{-n}{n-2}} - b'Y^{\frac{n+2}{n-2}} \leq 0, \tag{7.32}$$

Y is a constant subsolution if the opposite inequality is satisfied. Sufficient conditions for (7.32) to be satisfied are

$$k - r(\gamma) \leq b\theta^{\frac{4}{n-2}}Y^{\frac{4}{n-2}} \quad \text{and} \quad Y > 0, \tag{7.33}$$

together with (recall that $a \geq 0$ and $q \geq 0$)

$$\theta^{\frac{-3n+2}{n-2}}Y^{\frac{-4n+2}{n-2}} \leq 1, \quad \theta^{\frac{-n}{n-2}}Y^{\frac{-2n+2}{n-2}} \leq 1. \tag{7.34}$$

The inequality (7.33) is not obvious to satisfy when $b \geq 0$ but not $Inf_M b > 0$. If we can find a scalar function $k \geq 0$, $k \not\equiv 0$ and a number $\lambda > 0$ such that

$$k - r(\gamma) = \lambda^{\frac{4}{n-2}}b, \tag{7.35}$$

then (7.33) reads

$$\lambda \leq \theta Y, \tag{7.36}$$

and any number Y such that

$$Y \geq \lambda\theta^{-1} \tag{7.37}$$

satisfies the inequality (7.33). The question is now: can we satisfy (7.35) with $k \geq 0$, $k \not\equiv 0$, $\lambda > 0$? If the answer is yes the conformally transformed Equation (7.30) admits a constant supersolution m, a number such that

$$m \geq Max\{Sup\theta^{\frac{3n-2}{n-2}}, Sup\theta^{\frac{n}{n-2}}, Sup\lambda\theta^{-1}\}, \tag{7.38}$$

and, by the same proof a constant subsolution, a number ℓ such that:

$$0 < \ell \leq Min\{Inf\theta^{\frac{3n-2}{n-2}}, Inf\theta^{\frac{n}{n-2}}, Inf\lambda\theta^{-1}\} \leq m. \tag{7.39}$$

The answer to the question of the existence of λ and k depends on the SY (scalar-Yamabe) class of γ.

(1) If γ is in the positive SY class, with $r(\gamma) > 0$ and $b \geq 0$ any $\lambda > 0$ will give a positive k hence a solution $\theta > 0$ of 7.28, and from these λ and θ result sub and supersolutions

(2) If γ is in the zero SY class with $r(\gamma) = 0$ we can still have $k \geq 0$, $k \not\equiv 0$ and $\lambda > 0$ with $b \geq 0$, but we must now suppose $b \not\equiv 0$.

2. If γ is in the negative SY class, for instance γ chosen such that $r(\gamma) = -1$ the answer is obviously yes if $Inf_M b > 0$, since we can then choose $\lambda > (Inf_M b)^{-1}$, which implies $k = \lambda b - 1 > 0$. □

The following corollary replaces an inequality to be satisfied on M by an inequality on an open set of M.

Corollary 7.15 *In the case of the negative SY class the theorem still holds if* $\inf b > 0$ *only on a sufficiently large subset of* M.

Proof see Choquet-Bruhat (2004), Maxwell (2005), and Choquet-Bruhat *et al.* (2006). □

Remark 7.16 *Existence and non-existence results cover all cases when* $b \geq 0$ *is a constant*[50], *since then either* $\underset{M}{Inf} b > 0$ *or* $b \equiv 0$.

Remark 7.17 *If* $(M, \gamma, \bar{\psi})$ *is in the positive or zero SY class the constant supersolutions can be constructed directly for the given equation, with the same hypothesis on the coefficients, as follows*[51]. *For simplicity we write up the physical case* $n = 3$. *We denote by* $\varphi_0^4 \equiv y_0$ *the positive number which is the solution of the equation*

$$\underline{b} y_0^3 + \underline{r} y_0^2 - \underline{q} y_0 - \underline{a} = 0. \tag{7.40}$$

where $\underline{f}$ *denotes the mean value of a function* f *on* (M, γ):

$$\underline{f} \equiv \frac{1}{Vol(M, \gamma)} \int_M f \mu_\gamma. \tag{7.41}$$

[50] b is constant when τ is constant, as well as the unscaled matter sources and the potential for the scalar field, if such fields are present in the model.

[51] This is an extension of a method given by Moncrief in the case $n = 2$.

We define one function $v \in W_2^p$, with mean value zero on M, by solving the linear equation

$$\Delta_\gamma v = r\varphi_0 - a\varphi_0^{-7} - q\varphi_0^{-3} + b\varphi_0^5 \tag{7.42}$$

The function

$$\varphi_+ \equiv \varphi_0 + v - Inf v \geq \varphi_0, \quad \Delta_\gamma \varphi_+ \equiv \Delta_\gamma v, \tag{7.43}$$

is a supersolution if $r \geq 0$ because it holds that:

$$\Delta_\gamma \varphi_+ - f(., \varphi_+) = r(\varphi_0 - \varphi_+) - a(\varphi_0^{-7} - \varphi_+^{-7}) - q(\varphi_0^{-3} - \varphi_+^{-3}) + b(\varphi_0^5 - \varphi_+^5) \tag{7.44}$$

hence if $r \geq 0$

$$\Delta_\gamma \varphi_+ - f(., \varphi_+) \leq 0 \tag{7.45}$$

because $\varphi_+ \geq \varphi_0$.

In the zero Yamabe class the same type of argument holds, but we must use in addition the hypothesis $Inf_M b > 0$ to insure that $\varphi_0 > 0$.

7.3.2 *Case b of non-imposed sign*

We now consider the case where $b \equiv \frac{n-2}{4n}\tau^2 - q_0$ is not necessarily positive or zero. This may be the case in the presence of unscaled sources and/or of a positive scalar potential.

In the case where b may change sign we look for constant sub- and supersolutions. We set

$$f(., y) \equiv y^{-\frac{3n-2}{n-2}} h(., y). \tag{7.46}$$

where

$$h(., y) \equiv by^{\frac{4n}{n-2}} + ry^{\frac{4(n-1)}{n-2}} - a. \tag{7.47}$$

To treat a general n while avoiding unenlightening computations[52], we take $q \equiv 0$. The case $q \not\equiv 0$ can be studied by the same method. Constant sub- and supersolutions must satisfy on M the inequalities (see (7.25)

$$h(., \ell) \equiv b\ell^{\frac{4n}{n-2}} + r\ell^{\frac{4(n-1)}{n-2}} - a \leq 0 \tag{7.48}$$

$$h(., m) \equiv bm^{\frac{4n}{n-2}} + rm^{\frac{4(n-1)}{n-2}} - a \geq 0. \tag{7.49}$$

We suppose that M is the union of two closed sets, M_+ where $b \geq 0$ and M_- where $b \leq 0$. If (M, γ) is in the negative or zero Yamabe class with r of constant sign there is no number m satisfying (7.49) if M_- is not empty. We restrict here our study to the case where (M, γ) is in the positive scalar-Yamabe class,

$$\underset{M}{Inf} r > 0. \tag{7.50}$$

[52] In recent cosmological studies the case of b of changing sign occurs in the presence of scalar field with positive potential; other field sources are usually neglected.

We consider successively the domains M_+ and M_- and look for numbers ℓ_+ and m_+ [respectively ℓ_- and m_-] satisfying the inequalities (7.48) and (7.49) in M_+ [respectively M_-].

1. In $M_+ := M \cap \{b \geq 0\}$.

We use the trivial inequalities $rm^{\frac{4(n-1)}{n-2}} \geq r$ and $am^{\frac{2(n-1)}{n-2}} \geq a$ if $m \geq 1$, $a \geq 0, r \geq 0$ to find the sufficient conditions which follow, for ℓ_+ and m_+ to satisfy the inequalities (7.48) and (7.49)

$$m_+^{\frac{2(n-1)}{(n-2)}} \geq Max\left\{1, \frac{\underset{M_+}{Supa}}{\underset{M_+}{Inf}(b+r)}\right\}, \quad 0 < \ell_+ \leq Min\left\{1, \frac{\underset{M_+}{Infa}}{\underset{M_+}{Sup}(b+r)}\right\}. \quad (7.51)$$

2. In $M_- := M \cap b \leq 0$.

For each $x \in M_-$, $y \mapsto h(x,y)$ starts increasing from a value $y(x,0) = -a(x) \leq 0$, it will be possible to find a positive constant satisfying (7.48) on M_- if

$$\underset{M_-}{Inf}\ a > 0, \quad (7.52)$$

$y \mapsto h(x,y)$ decreases after its maximum attained at

$$[y_{Max}(x)]^{\frac{4}{n-2}} = \frac{(n-1)r(x)}{n|b(x)|} \quad (7.53)$$

Some calculus shows that, on M_-,

$$h(.,y_{Max}) \equiv \frac{1}{n}\left[\frac{(n-1)}{n}\right]^{n-1} \frac{r^n}{|b|^{n-1}} - a. \quad (7.54)$$

Therefore the maximum $h(.,y_{Max})$ is non-negative if and only if

$$n((n-1)^2|b|^{n-1}a \leq r^n. \quad (7.55)$$

Remark 7.18 *This condition corresponds in the case $n = 3$ to the well-known condition for the third-order polynomial h in y^4 to have three real roots,*

$$27b^2a < 4r^3. \quad (7.56)$$

For general n, when $y_{Max}(x) \geq 0$ (and (7.58 holds) the equation $h(x,y) = 0$ has two positive roots $0 < z_1(x) \leq z_2(x)$. The numbers $X(x)$ and $Y(x)$ satisfy the inequalities (7.48) and (7.49) at x if

$$0 < X(x) \leq z_1(x) \leq Y(x) \leq z_2(x). \quad (7.57)$$

There exist constant numbers ℓ_- and m_- satisfying (7.48) and (7.49) on M_- if, in addition to (7.57) and (7.55), it holds that:

$$\underset{M_-}{Supz_1(x)}] \leq \underset{M_-}{Inf} z_2(x) \quad (7.58)$$

The following existence theorem is now a consequence of the general existence Theorem 7.13.

Theorem 7.19 *The Lichnerowicz equation on a compact manifold (M, γ) in the positive Yamabe class, with $r, a, b \in L^\infty$, admits a solution $\varphi > 0$ if the coefficients are such that the conditions for the existence of ℓ_+, ℓ_-, m_+ and m_- are satisfied and if*

$$Max(\ell_+, \ell_-) \leq Min(m_+, m_-). \tag{7.59}$$

Theorem 7.19 is not optimal. It only gives sufficient conditions for the existence of a solution, and leaves space for further research if physical problems motivate it, in particular for existence in negative or zero Yamabe classes.

A recent paper by Hebey *et al.*[53] introduces a variational method to solve the Lichnerowicz equation which does not use super- and subsolutions. They obtain an existence theorem with a simpler formulation of the inequalities to be satisfied by the coefficients, but they suppose that the operator $\Delta_\gamma - r$ is coercive, that is, again, that γ is in the positive Yamabe class. The existence of solutions when b takes negative values and (M, γ) is in the negative or the zero Yamabe class is still open.

8 System of constraints

Recall that, in all cases, solutions $(\bar{g}, K)$ of the original constraints are deduced from solutions (φ, X) of the CF constraints with data γ, τ, U by the formulae

$$\bar{g}_{ij} = \varphi^{\frac{4}{n-2}} \gamma_{ij}, \qquad K^{ij} = \varphi^{-2(n+2)/(n-2)} [(\mathcal{L}_{\gamma,conf} X)^{ij} + U^{ij}] + \frac{1}{n} \bar{g}^{ij} \tau, \tag{8.1}$$

where U is the traceless "radiation data" symmetric 2-tensor introduced for the solution of the momentum constraint.

8.1 *Constant mean curvature τ, sources with York-scaled momentum*

When τ is a constant the conformally formulated momentum constraint with York scaled sources does not depend on φ. We can therefore solve this linear system for X and insert afterwards the result in the semilinear equation for φ. While the metric γ and the coefficients b and q are given quantities, the function a depends, in addition to the specified tensor U, on the solution X of the momentum constraint. We state an existence theorem which summarizes results of the previous sections. We use the definitions given in Section 3.4 in the case where the sources of the Hamiltonian constraint are also York scaled. We leave to the reader the formulations of variants of the following existence theorem in the case where these sources are not scaled, for instance in the presence of a scalar field with positive potential, using Theorem 7.19.

[53] Hebey, E., Pacard, F., and Pollack, D. 2008, to appear.

Theorem 8.1 *Let $\gamma \in M_2^p$, $p > \frac{n}{2}$. Let $q \in L^p$ (definition 7.22); Let $\tilde{J}_1 \in L^p$, L^2-orthogonal to the conformal Killing fields of (M, γ), be given as well as $U \in W_1^p$, and $\tau =$constant. Suppose the momentum is York-scaled ($J_0 \equiv 0$). Then the CF scalar-Einstein constraints admit a solution $X, \varphi \in W_2^p$, $\varphi > 0$, in the following cases.*

1. *Case $b := \frac{n-1}{n}\tau^2 - q_0 \geq 0$ on M, condition satisfied in particular when unscaled matter sources and negative or zero potential of the scalar field are absent.*

   (a) *(M, γ) is in the positive Yamabe class $r > 0$ and $|\tilde{J}|_1 + q \not\equiv 0$ on M.*

   (b) *(M, γ) is in the zero Yamabe class $r = 0$, $\tau \neq 0$ and $|J|_1 + q \not\equiv 0$ on M.*

   (c) *(M, γ) is in the negative Yamabe class, $p > n$, $q \in L^\infty$, $\tau \neq 0$.*

   The solution is unique, except if (M, γ) is in the zero Yamabe class and $|J|_1 + |U| + q + \tau^2 \equiv 0$ (in this case $\tilde{K} = 0$ and φ is any positive constant).

2. *Case where b is not positive everywhere on M. (M, γ) in the positive Yamabe class, $p > n$, and the hypotheses of Theorem 7.19 are satisfied.*

Proof Under the hypotheses the momentum constraint (6.1) has a solution $X \in W_2^p$, since the vector F given by (6.2) is here independent of φ. This satisfies the hypothesis of Theorem 6.2. The tensor $\tilde{K} \equiv \mathcal{L}_{\gamma,conf} X + U$ is therefore in W_1^p. The function $a \equiv k_n(|\tilde{K}|_\gamma^2 + \tilde{\pi}^2)$ is also in L^p as soon as $p > \frac{n}{2}$.

1(a) and (b). The existence Theorem 7.14, positive or zero Yamabe class, applies because if $\tilde{J}_1 \not\equiv 0$ the equation

$$D_i \tilde{K}^{ij} = J_1^j$$

implies $\tilde{K} \not\equiv 0$; hence $a \not\equiv 0$. Therefore $|\tilde{J}_1| + q \not\equiv 0$ implies $a + q \not\equiv 0$.

1(c) and 2. If $p > n$ the tensor $\tilde{K}$ is in C^0, hence the function a is also in $C^0 \subset L^\infty$. We apply in case 1(c) Theorem 7.14; in case 2 we apply Theorem 7.19. □

8.2 *Solutions with $\tau \not\equiv$ constant or $J_0 \not\equiv 0$*

We will show that there exists a whole neighbourhood of low regularity solutions with non-constant[54] τ, or sources with unscaled momentum, near a low regularity solution with constant τ and York-scaled sources.

Lemma 8.2 *The mapping defined by:*

$$\Phi : (x, y) \mapsto \Phi(x, y) \equiv (\mathcal{H}(x, y), \mathcal{M}(x, y)), \tag{8.2}$$

$$x \equiv (\gamma, \tau, U, \bar{\psi}, \tilde{\pi}, q_1, q_2, q_0, J_1, J_0), \quad y \equiv (\varphi, X) \tag{8.3}$$

[54] There are cosmological spacetimes which do not admit surfaces of constant τ (Bartnik (1988), case of dust sources; Chruściel, Isenberg and Pollack, vacuum case).

$$\mathcal{H}(x,y) \equiv \Delta_\gamma \varphi - r(\gamma,\bar{\psi})\varphi + a(U,\tilde{\pi},X)\varphi^{-\frac{3n-2}{n-2}} + q_1\varphi^{-\frac{n}{n-2}} + q_2\varphi^{-\frac{6-n}{n-2}} - \left(q_0 - \frac{n-2}{4n}\tau^2\right)\varphi^{\frac{n+2}{n-2}} \tag{8.4}$$

$$\mathcal{M}^i(x,y) \equiv (\Delta_{\gamma,conf}X)^i - \left\{\frac{n-1}{n}\varphi^{2n/(n-2)}\partial^i\tau + J_1^i + J_0^i\varphi^{2(n+2)/(n-2)} - D_jB^{ij}\right\} \tag{8.5}$$

is a differentiable map into the Banach space $L^p \times^1 \otimes L^p$, *if* $p > \frac{n}{2}$, *from the open set* $\Omega \equiv \Omega_1 \times \Omega_2$ *of the Banach space* $\mathbf{B} \equiv \mathbf{B}_1 \times \mathbf{B}_2$, *defined by*[55]

$$\Omega_1 \equiv \mathbf{B}_1 \cap \{\gamma \in M_2^p\}, \quad \Omega_2 \equiv \mathbf{B}_2 \cap \{\varphi > 0\} \tag{8.6}$$

$$\mathbf{B}_1 \equiv (W_2^p \times W_1^p \times ({}^2\otimes W_1^p) \times L^p \times L^p \times^1 \otimes L^p \times^1 \otimes L^p), \tag{8.7}$$

$$\mathbf{B}_2 \equiv W_2^p \times ({}^1\otimes W_2^p). \tag{8.8}$$

Proof The mapping Φ is a differentiable mapping from Ω into $L^p \times^1 \otimes L^p$ because the linear mapping obtained by linearization

$$\delta\Phi_{x,y} : (\delta x, \delta y) \mapsto \delta\Phi_{x,y}(\delta x, \delta y) \equiv (\delta\mathcal{H}_{x,y}(\delta x,\delta y), \delta\mathcal{M}_{x,y}(\delta x,\delta y)) \tag{8.9}$$

is a mapping from the Banach space $\mathbf{B}$ into the Banach space $L^p \times^1 \otimes L^p$ for each $(x,y) \in \Omega$, as can be checked by the embeddings and multiplication properties of Sobolev spaces.

The partial derivative with respect to y is given by

$$\delta\Phi_y : \delta y \mapsto \delta\Phi_y(\delta y) \equiv (\delta\mathcal{H}_y(\delta y), \delta\mathcal{M}_y(\delta y)) \tag{8.10}$$

with

$$\delta\mathcal{H}_y(\delta y) \equiv \Delta_\gamma\delta\varphi - \left\{r + \frac{3n-2}{n-2}a\varphi^{-\frac{2n}{n-2}} + \frac{n}{n-2}q_1\varphi^{-\frac{2(n-1)}{n-2}} + \frac{6-n}{n-2}q_2\varphi^{-\frac{4}{n-2}} + \frac{n+2}{n-2}(\bar{b})\varphi^{\frac{4}{n-2}}\right\}\delta\varphi + a'_X\varphi^{\frac{-3n+2}{n-2}}\delta X \tag{8.11}$$

and

$$\delta\mathcal{M}_{x,y}(\delta x) \equiv (\Delta_{\gamma,conf}\delta X)^i - \left(\frac{2(n-1)}{n-2}\varphi^{\frac{n+2}{n-2}}\partial^i\tau + \frac{2(n+2)}{n-2}\varphi^{\frac{n+6}{n-2}}J_0^i\right)\delta\varphi.$$

□

Theorem 8.3 *Let* $(\bar{g},\bar{K})$, $\bar{K} = \bar{U} + \mathcal{L}_{\bar{g},conf}\bar{X} + \frac{\bar{g}}{n}\bar{\tau}^2$ *be a solution of the scalar-Einstein constraints on the compact n-manifold M, with* $Tr\bar{K} \equiv \bar{\tau}$ *a constant,*

[55] To help the reader follow our reasoning we denote by ${}^m\otimes W_s^p$ the space of W_s^p m-tensor fields.

$\bar{g} \in M^p_2$, $\bar{X} \in W^p_2$, $\bar{U} \in W^p_1$, $p > \frac{n}{2}$, York-scaled sources $\bar{q}_1, \bar{q}_2$, $\bar{q}_0$, $\bar{J}_1 \in L^p$. Suppose that sources which are not York-scaled are such that $\bar{J}_0 \equiv 0$, and $\bar{b} \geq 0$. Suppose $(M, \bar{g})$ admits no conformal Killing field. Then, there is a neighbourhood of $\bar{\tau}, \bar{U}$, $\bar{q}_1$, $\bar{q}_2$, $\bar{J}_1$, $\bar{J}_0$, $\bar{b}$ in $\mathbf{B}_1$ such that the scalar-Einstein constraints with non-constant τ and [or] $J_0 \not\equiv 0$ admit a solution (g, K) in a $W^p_2 \times W^p_1$ neighbourhood of $(\bar{g}, \bar{K})$, if $\bar{a}$, $\bar{q}_1$, and $\bar{b}$ are not all identically zero.

Proof The function $\bar{\varphi} = 1$ together with $\bar{X} \in W^p_2$ is a solution of the CF constraints with the given data, $\bar{\gamma} \equiv \bar{g} \in M^p_2$, $\bar{\tau}$, $\bar{U}$, etc. We will have proved the existence of a solution $\varphi \in W^p_2$, $\varphi > 0$, when $\gamma \in M^p_2$, $\tau \in W^p_1$, $U \in W^p_1$, q_1, q_2, q_0, J_1, J_0, b are in a sufficiently small neighbourhood in the Banach space $\mathbf{B}$ of their overlined values if the partial derivative of Φ with respect to y at a point $(\bar{x}, \bar{y})$ is an isomorphism of Banach spaces. At a point where $\partial_i \bar{\tau} \equiv 0$ and $\bar{J}_0 \equiv 0$ the condition is for the linear system

$$(\Delta_{\bar{g},conf} \delta X)^i = Y \tag{8.12}$$

and (recall that $\bar{\varphi} = 1$)

$$\Delta_{\bar{g}} \delta\varphi - \left\{ \bar{r} + \frac{3n-2}{n-2}\bar{a} + \frac{n}{n-2}\bar{q}_1 + \frac{6-n}{n-2}\bar{q}_2 + \frac{n+2}{n-2}\bar{b} \right\} \delta\varphi + \bar{a}'_X \delta X = \psi \tag{8.13}$$

to have one and only one solution $\delta X \in^1 \otimes W^p_2$, $\delta\varphi \in W^p_2$ for each $Y \in^1 \otimes L^p$, $\psi \in L^p$. This result holds for (8.12) if $\bar{g}$ admits no conformal Killing field. It holds then for (8.13) if the coefficient of $\delta\varphi$, which belongs to L^p, $p > \frac{n}{2}$ by hypothesis, is positive or zero, and not identically zero. It is obviously so when $\bar{r} > 0$, $n \leq 6$, $\bar{b} \geq 0$. To treat general values of $\bar{r}$ we remark, as O'Murchada and York (1974)[56] in their study of linearization stability, that since $(\bar{g}, \bar{K})$ satisfies the Hamiltonian constraint, hence the Lichnerowicz equation with $\bar{\varphi} = 1$, it holds that:

$$-\bar{r} + \bar{a} + \bar{q}_1 + \bar{q}_2 - \bar{b} = 0. \tag{8.14}$$

The linear equation (8.13) reads therefore

$$\Delta_{\bar{g}} \delta\varphi - \left\{ 2\bar{a} + \frac{2(n-1)}{n-2}\bar{q}_1 + \frac{4}{n-2}\bar{q}_2 + \frac{4}{n-2}\bar{b} \right\} \delta\varphi + \bar{a}'_X \delta X = \psi \tag{8.15}$$

The coefficient of the order term $\delta\varphi$ is positive or zero, and not identically zero, if $\bar{a}$, $\bar{q}_1$, $\bar{q}_2$, and $\bar{b}$ are not all identically zero; the equation has then one and only one solution $\delta\varphi$. □

The result does not extend in a straightforward way[57] to the case where $\bar{g}$ admits conformal Killing fields because $\mathcal{M}$ does not takes its values in the space of vector fields orthogonal to such fields. Holst, Nagy and Totgenel (2008) obtain solutions with small rediation data but large $\partial\tau$, positive Yamabe and $9 > 0$.

[56] O'Murchada, N. and York, J. (1974) *Phys. Rev. D*, **10**, 428–46.
[57] An analogous problem occurs in linearization stability.

9 Solutions on asymptotically Euclidean Manifolds

We treat now the case where (M, γ) is asymptotically Euclidean[58], and $n \geq 3$[59]. Solutions on asymptotically Euclidean manifolds are important to model the motion of isolated bodies; they were the first to be found in special cases.

A Riemannian manifold (M, γ) is called **asymptotically Euclidean** (see Appendix I) if there exists a smooth Riemannian manifold (M, e) Euclidean at infinity[60] and γ tends to e at infinity in each end. In one end U, and in terms of the canonical coordinates x^i of E^n, the metric e has components $e_{ij} = \delta_{ij}$. The metric γ tends to e at infinity if in these coordinates $\gamma_{ij} - \delta_{ij}$ tends to zero. To solve elliptic systems on asymptotically Euclidean manifolds we used weighted Sobolev spaces (see Appendix II). We could use also weighted Hölder spaces[61], but they are not well adapted to the related evolution problems.

A **weighted Sobolev space** $W^p_{s,\delta}$, $1 \leq p \leq \infty$, s a positive or zero integer, δ a real number, of tensors of some given type on the manifold (M, e) Euclidean at infinity is the space of tensors of some given type which admit generalized e-covariant derivatives of order up to s for which the following norm is finite:

$$\|u\|_{W^p_{s,\delta}} = \left\{ \sum_{0 \leq m \leq s} \int_V | \partial^m u |^p \, (1 + d^2)^{\frac{1}{2} p(\delta + m)} d\mu \right\}^{1/p}, \tag{9.1}$$

where ∂, $|\ |$, and $d\mu$ denote the covariant derivative, norm, and volume element in the metric e, and d is the distance in the metric e from a point of M to a fixed point. If (M, e) is a Euclidean space one can choose $d = r := \sum_i \{(x^i)^2\}^{\frac{1}{2}}$, the Euclidean distance to the origin. The space $\mathcal{D}$ of C^∞ tensors with compact support is dense in $W^p_{s,\delta}$ regardless of s and δ if $p < \infty$.

If s and δ are large enough, a function (or tensor field) in $W^p_{s,\delta}$ is continuous and tends to zero at infinity. Namely if we define C^m_β to be the Banach space of weighted C^m functions (or tensor fields) on (M, e) with norm given by

$$\|u\|_{C^m_\beta} \equiv \sum_{0 \leq \ell \leq m} \sup_M (|\partial^\ell u| (1 + d^2)^{\frac{1}{2}(\beta + \ell)}).$$

Recall (see Appendix I) the following inequality, where C is a number depending only on (M, e),

$$\|u\|_{C^m_\beta} \leq C ||u||_{W^p_{s,\delta}}, \quad \text{if} \quad s > m + \frac{n}{p}, \quad \delta > \beta - \frac{n}{p}.$$

We see that $u \in W^p_{s,\delta}$ implies that u is continuous and tends to zero at infinity if $s > \frac{n}{p}$ and $\delta > -\frac{n}{p}$.

[58] See Choquet-Bruhat, Y., Isenberg, J., and York, J. (2000) *Phys. Rev. D*, **3**, 61.

[59] For definitions of energy and momentum in $2+1$ gravity, see Ashtekar, A. and Dryer, O. (2003) gr-qc 0206024 V3.

[60] (See Appendix I).

[61] Chaljub-Simon, A-ànd Choquet-Bruhat, Y. (1979) *Ann. Univ. Toulouse*, **1**, 25–39.

Let (M, e) be a manifold Euclidean at infinity. The Riemannian manifold (M, γ) is said to be (p, σ, ρ) asymptotically Euclidean if $\gamma - e \in W^p_{\sigma,\rho}$. If $\gamma - e \in W^p_{\sigma,\rho}$ with $\sigma > \frac{n}{p}$, $\rho > -\frac{n}{p}$, then γ is C^0 and $\gamma - e$ tends to zero at infinity. The set of (p, σ, ρ) asymptotically Euclidean metrics (i.e. positive definite symmetric 2-tensors) is denoted $M^p_{\sigma,\rho}$; it is an open set of $W^p_{\sigma,\rho}$, $\sigma > \frac{n}{p}$, $\rho > -\frac{n}{p}$.

We recall the multiplication lemma

$$W^p_{s_1,\delta_1} \times W^p_{s_{2,}\delta_2} \subset W^p_{s,\delta} \quad \text{if} \quad s < s_1 + s_2 - \frac{n}{p}, \quad \delta < \delta_1 + \delta_2 + \frac{1}{p}. \tag{9.2}$$

10 Momentum constraint

Recall that the momentum constraint reads

$$D_j(\mathcal{L}X)^{ij} \equiv (\Delta_{\gamma,conf} X)^i = F^i(., \varphi) \tag{10.1}$$

with

$$F^i(., \varphi) \equiv D_j U^{ij} + \frac{n-1}{n} \varphi^{2n/(n-2)} \gamma^{ij} \partial_j \tau + \varphi^{2(n+2)/(n-2)} J_0^i + J_1^i \tag{10.2}$$

where τ is a given function on M and U^{ij} is a given symmetric traceless tensor field. The scaled and unscaled sources J_1 and J_0 are also considered as known. We suppose momentarily that φ is also a known function; in fact it disappears from the equation if $\partial\tau \equiv 0$ and $J_0 \equiv 0$.

The following theorem is a direct consequence of the isomorphism theorem 4.7 proved in Appendix II for the conformal Laplace operator $\Delta_{\gamma,conf}$, because the hypotheses made below imply, due to the Sobolev embedding and multiplication properties, that $F(., \varphi) \in W^p_{0,\delta+2}$.

Theorem 10.1 *Let (M, γ) be a $W^p_{2,\delta}$ asymptotically Euclidean manifold, $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$. Let U, $\tau \in W^p_{1,\delta+1}$ and $J_1, J_0 \in W^p_{0,\delta+2}$ be given. Suppose also that φ is known, with $\varphi > 0$, $(1-\varphi) \in W^p_{2,\delta}$. Then the momentum constraint has one and only one solution $X \in W^p_{2,\delta}$ if in addition $\delta < n - 2 - \frac{n}{p}$.*

If $\partial\tau$ and J_0 are zero the conditions on φ are irrelevant.

The proof of the following corollary rests on corollary 4.8, Appendix II.

Corollary 10.2 *If in addition $\gamma \in H_{2+s,\delta}$, $U, \tau \in H_{s+1,\delta+1}$ and $J_1, J_0 \in H_{s,\delta+2}$, and $(1-\varphi) \in H_{s+2,\delta}$, $-\frac{n}{2} < \delta < \frac{n}{2} - 2$, then the solution X belongs to $H_{s+2,\delta}$.*

Remark 10.3 *In the conformal thin sandwich formulation one must make, in addition, the hypothesis $\tilde{N} = 1 + \nu > 0$, $\nu \in W^p_{2,\delta}$ for the theorem, $\nu \in H_{s+2,\delta}$ for the corollary.*

11 Solution of the Lichnerowicz equation

We have called the Lichnerowicz equation an equation of the form

$$\mathcal{H}(x,\varphi) \equiv \Delta_\gamma \varphi - f(x,\varphi) = 0, \tag{11.1}$$

with

$$f(x,\varphi) \equiv r\varphi - a\varphi^{-\frac{3n-2}{n-2}} - q_1\varphi^{-\frac{n}{n-2}} - q_2\varphi^{-\frac{6-n}{n-2}} + b\varphi^{\frac{n+2}{n-2}},$$

where the coefficients r, a, b, q_1, and q_2 are given by (3.47) and (3.48).

We first prove the following easy lemma.

Lemma 11.1

1. *If* (M,γ) *is a* $(p,2,\delta)$ *asymptotically Euclidean manifold with* $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$, *then* $R(\gamma) \in W^p_{0,\delta+2}$.
2. *If in addition* $\gamma \in H_{s,\delta}$, $s > \frac{n}{2}$, $s \geq 2$, $\delta > -\frac{n}{2}$, *then* $R(\gamma) \in H_{s-2,\delta+2}$.

Proof $R(\gamma)$ is a sum of terms of the form $\gamma\partial^2\gamma$, and $\gamma\partial\gamma\partial\gamma$, with $\gamma - e \in W^p_{2,\delta}$, $\partial\gamma \in W^p_{1,\delta+1}$, $\partial^2\gamma \in W^p_{0,\delta+2}$. Under the hypothesis made on p, and δ and σ, the Sobolev embedding theorem shows that $\gamma - e$ is continuous and bounded on M; the multiplication lemma, Appendix I, completes the proof. An analogous proof holds for 2. □

11.1 *Uniqueness theorem*

The following theorem is analogous to Theorem 7.12 proved in the case of a compact M.

Theorem 11.2 *(uniqueness) The Lichnerowicz equation, with* $f(.,\varphi)$ *given by (3.47), on* (M,γ), $\gamma \in M^p_{2,\delta}$, $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$ *has at most one positive solution* φ, $\varphi - 1 \in W^p_{2,\delta}$, *if* $r, a, q_i, b \in W^p_{0,\delta+2}$, $a \geq 0$, $q_1 \geq 0$, $(6-n)q_2 \geq 0$, $b \geq 0$.

Proof One finds as in the proof of Theorem 7.12 that $u \equiv \varphi_1\varphi_2^{-1} - 1$ satisfies an equation of the form

$$\Delta_{\varphi_2^{2p}\gamma} u - \lambda u = 0 \tag{11.2}$$

with $\lambda \geq 0$. This inequality is a sufficient condition, with $u \in W^p_{2,\delta}$, to ensure that $u \equiv 0$. □

11.2 *Generalized Brill–Cantor theorem*

There is no theorem for asymptotically Euclidean manifolds analogous to the Yamabe theorem which proves the existence of a conformal metric with constant scalar curvature for any Riemannian metric on a compact manifold. However, an interesting theorem was proved by Brill and Cantor (1981)[62], and can be generalized in the presence of a scalar field as follows.

[62] Under a somewhat more restrictive hypothesis on regularity; Brill, D. and Cantor, M. (1981) *Compos. Math.*, **43**, 317–25. Also it was pointed out by Maxwell, D. (2003) gr-qc

Theorem 11.3 *Let (M,γ) be a $(p,2,\delta)$ asymptotically Euclidean manifold with $p>\frac{n}{2}$, $\delta>-\frac{n}{p}$, and let $\bar{\psi}\in W^p_{2,\delta}$ be a scalar field on M. There exists on M a $(p,2,\delta)$ asymptotically Euclidean metric γ' conformal to γ such that $r(\gamma',\bar{\psi})=0$ if and only if $(M,\gamma,\bar{\psi})$ satisfy the following inequality*

$$\int_M \left\{|Df|^2+r(\gamma,\bar{\psi})f^2\right\}\mu_\gamma>0 \tag{11.3}$$

for every function f on M with $f\in W^p_{2,\tilde{\delta}}$, $\tilde{\delta}>-\frac{n}{p}+\frac{n}{2}-1$ ($\tilde{\delta}\geq -1$ if $p=2$), $f\not\equiv 0$.

Proof (M,γ) is conformal to $(M,\gamma')\in M^p_{2,\delta}$ with $r(\gamma',\bar{\psi})\equiv 0$ if and only if there exists a function $\varphi>0$, such that $\gamma'=\varphi^{4/(n-2)}\gamma\in M^p_{2,\delta}$ and

$$\triangle_\gamma\varphi-r(\gamma,\bar{\psi})\varphi=0. \tag{11.4}$$

That is setting $\varphi\equiv 1+u$:

$$\Delta_\gamma u-r(\gamma,\bar{\psi})u=r(\gamma,\bar{\psi}). \tag{11.5}$$

1. Suppose that the condition (11.3) is satisfied. Equation (11.5) is linear elliptic, with $\Delta_\gamma-r(\gamma,\bar{\psi})$ an injective operator on $W^p_{2,\delta}$, and satisfies the hypothesis of Theorem 4.4 of Appendix II; therefore it admits a solution $u\in W^p_{2,\delta}\subset C^0_\alpha$. It remains to prove that $\varphi\equiv 1+u$ is positive; then $\varphi^{4/(n-2)}\gamma\in M^p_{2,\delta}$. One cannot use directly the maximum principle because $r(\gamma,\bar{\psi})$ is not necessarily positive. Inspired by Brill and Cantor we consider the family of equations

$$\triangle_\gamma\varphi-kr(\gamma,\bar{\psi})\varphi=0 \quad \text{i.e.} \quad \triangle_\gamma u-kr(\gamma,\bar{\psi})u=kr(\gamma,\bar{\psi}) \tag{11.6}$$

with $k\in[0,1]$ a number. Each of these equations admits a solution $u_k\in W^p_{2,\delta}\subset C^0_\alpha$, and the C^0_α norm of u_k depends continuously on k. The set $S=:\{u_k\in C^0_\alpha, u_k>-1\}$ is open in C^0_α and non-empty because for $k=0$ it holds that $u_0=0$ (i.e. $\varphi_0=1$). To show that it is closed suppose that $u_{k'}$ belongs to its boundary ∂S; then $u_{k'}\geq -1$, $\varphi_{k'}\geq 0$. Suppose that $\varphi_{k'}$, a solution of the elliptic equation (11.6), vanishes at a point of M; then by the weak Harnack inequality (Trudinger, 1973) there is a ball B_R of centre x and a number C such that

$$||\varphi_{k'}||_{L^q(B_{2R})}\leq CInf_{B_R}\varphi_{k'}=0; \tag{11.7}$$

0307117, that the original definition of Brill–Cantor, carried over in Choquet-Bruhat *et al.*, which used compactly supported f, was incorrect, because the limit of positive functions is only known to be non-negative. Maxwell replaces it by

$$Inf_{f\in\mathcal{D},f\not\equiv 0}\left(\int_M\left\{|Df|^2+k_nR(\gamma)f^2\right\}\mu_\gamma\right)/||f||^2_{L^{2n/(n-2)}}>0$$

Our definition, which generalizes the formula (4.1) of Appendix II, used for proving injectivity, also remedies the original defect.

hence $\varphi_{k'} = 0$ in B_R and also, by continuity, on M. This is impossible because $\varphi_{k'}$ tends to 1 at infinity. Hence $\varphi_{k'} > 0$. The subset S of C^0_α, being both open and closed, covers this whole functional space.

2. Conversely, suppose that $\varphi > 0$ exists and solves the equation satisfying the hypothesis of the theorem. Then we will show that for any $f \not\equiv 0$, $f \in W^p_{2,\delta}$ the inequality (11.3) holds. We set $\theta = f\varphi^{-1}$, then $\theta \in W^p_{2,\delta} \subset C^0_\alpha$. We have by elementary calculus:

$$|Df|^2 = |D\theta|^2\varphi^2 + \varphi D\varphi.D(\theta^2) + \theta^2|D\varphi|^2 \ . \tag{11.8}$$

The following integration by parts holds for the considered functions:

$$\int_M \varphi D\varphi.D(\theta^2)\mu_\gamma = \int_M -\theta^2 D(\varphi D\varphi)\mu_\gamma \ , \tag{11.9}$$

therefore,

$$\int_M \varphi D\varphi.D(\theta^2)\mu_\gamma = \int_M -\theta^2(\varphi\triangle_\gamma\varphi + |D\varphi|^2)\mu_\gamma \tag{11.10}$$

and

$$\int_M |Df|^2\mu_\gamma = \int_M \{|D\theta|^2\varphi^2 - \theta^2\varphi\triangle_\gamma\varphi\}\mu_\gamma \ . \tag{11.11}$$

We have $D\theta \not\equiv 0$ since $\theta \in C^0_\alpha$ tends to zero at infinity and cannot be a constant without being identically zero, impossible if $f \not\equiv 0$. Hence when $\varphi > 0$ satisfies Equation (11.4) the function $f \in W^p_{2,\delta}$, $f \not\equiv 0$ satisfies the inequality:

$$\int_M \left\{|Df|^2 + r(\gamma, \bar\psi)f^2\right\}\mu_\gamma > 0. \tag{11.12}$$

□

Sets $(M, \gamma, \bar\psi)$ satisfying (11.12) are said to be in the positive **SY class**. The name can be misleading, as shown by the following remark.

Remark 11.4 *The same kind of proof as those used above shows that, under the same hypothesis, there exists on M a metric γ' conformal to γ such that $r(\gamma', \bar\psi) < 0$.*

Remark 11.5 *If the initial manifold is maximal (necessarily so in the asymptotically Euclidean case if the initial data are such that $\tau := TrK = $ constant), the physical metric $\bar g$ that solves the constraints together with the symmetric two-tensor K has necessarily a non-negative scalar curvature $R(\bar g)$ if the sources have positive or zero energy. That is, $R(\bar g) \geq 0$, with $R(\bar g) \not\equiv 0$ except in a vacuum and for $K \equiv 0$, i.e. an "instant of time symmetry". Therefore, the physical metric $\bar g$ on an initial maximal submanifold is in the positive Yamabe class and all metrics γ used as substrata to obtain it must be in that class. This result does not necessarily hold for the scalar-Yamabe class.*

11.3 *Existence theorems*

To prove the existence of solutions by a constructive method we will apply Theorem 5.3 of Appendix II, which uses sub- and supersolutions.

Constant sub- and super solutions are not natural in the asymptotically Euclidean case. We will eventually introduce some intermediate steps to obtain non-constant sub- and supersolutions.

To avoid cumbersome computations in the case $n \geq 6$, we set $q = q_1 + q_2$ and suppose $q_2 \equiv 0$ if $n \geq 6$. Recall that $a \geq 0$, $q \geq 0$.

The following theorem applies in particular when $q_0 \equiv 0$, that is when there are no unscaled matter sources and the potential of the scalar field is such that $V(\bar{\psi}) \leq 0$.

Theorem 11.6 *Let (M, γ) be a $M^p_{2,\delta}$ manifold, $p > \frac{n}{2}$. Let $\bar{\psi}$ be a scalar field on M with potential $V(\bar{\psi})$, such that $\bar{\psi} \in W^p_{2,\delta}$ and $V(\bar{\psi}) \in W^p_{0,\delta+2}$. Suppose that $(M, \gamma, \bar{\psi})$ is in the positive SY class and $b \geq 0$. The Lichnerowicz equation*

$$\Delta_\gamma \varphi - r\varphi + a\varphi^{-\frac{3n-2}{n-2}} + q\varphi^{\frac{-n}{n-2}} - b\varphi^{\frac{n+2}{n-2}} = 0, \tag{11.13}$$

$$a, b, q \in W^p_{0,\delta+2}, \quad \delta > -1 + \frac{n}{2} - \frac{n}{p},\ \delta \geq -1 \quad \text{if } p = 2, \tag{11.14}$$

has one and only one solution, $\varphi = 1+u$, $u \in W^p_{2,\delta}$, if in addition $\delta < -2+n-\frac{n}{p}$. It can be obtained by iterations.

Corollary 11.7 *If moreover $\gamma \in M_{2+s,\delta}$ and $a, b \in H_{s,\delta+2}$, then $u \in H_{s+2,\delta}$.*

Proof **Uniqueness**: results from the general Theorem 11.2. It can also be proved directly using the monotonicity of the non-linear term.

Existence. We consider the conformally transformed equation to a metric such that $r(\gamma, \bar{\psi}) = 0$:

$$\Delta_\gamma \varphi + a\varphi^{-\frac{3n-2}{n-2}} + q\varphi^{-\frac{n}{n-2}} - b\varphi^{\frac{n+2}{n-2}} = 0, \tag{11.15}$$

1. We first consider the equation with $b \equiv 0$:

$$\Delta_\gamma \varphi + a\varphi^{-\frac{3n-2}{n-2}} + q\varphi^{-\frac{n}{n-2}} = 0 \tag{11.16}$$

This equation admits a constant subsolution $\varphi_- = 1$ but no finite constant supersolution. However, it admits a non-constant supersolution, namely the function $\varphi_+ = 1 + u_+$ with $u_+ \in W^p_{2,\delta}$ solution of the linear equation

$$\Delta_\gamma u_+ = -(a + q). \tag{11.17}$$

Indeed the maximum principle shows that $u_+ \geq 0$; hence $\varphi_+ \geq 1$ and

$$\Delta_\gamma \varphi_+ = -(a + q) \leq -(a\varphi_+^{-\frac{3n-2}{n-2}} + q\varphi_+^{-\frac{n}{n-2}}). \tag{11.18}$$

We can apply the general theorem to prove the existence of a solution φ_1.

2. We consider the equation with $a \equiv q \equiv 0$:

$$\Delta_\gamma \varphi - b\varphi^{\frac{n+2}{n-2}} = 0. \tag{11.19}$$

This equation admits the subsolution $\varphi_- = 0$ and the supersolution $\varphi_+ = 1$. It admits therefore a solution φ_2, with $1 - \varphi_2 \in W^p_{2,\delta}$, and $0 \leq \varphi_2 \leq 1$. We prove that $\varphi_2 > 0$ by an argument similar to the one used in the proof of the Brill–Cantor theorem: we consider the family of equations

$$\triangle_\gamma \varphi - kb\varphi^{\frac{n+2}{n-2}} = 0 \tag{11.20}$$

with $k \in [0,1]$ a number. Each of these equations admits one solution $\varphi_k = 1 + u_k \geq 0$, with $u_k \in W^p_{2,\delta} \subset C^0_\alpha$, and the C^0_α norm of u_k depends continuously on k. The proof continues as in the Brill–Cantor Theorem 11.3.

3. Consider the general equation (11.20). It admits φ_1 as a supersolution and φ_2 as a subsolution.

Proof of corollary. See Corollary 5.4 in Appendix II. □

Remark 11.8 *It is essential for physical applications to admit isolated sources, hence discontinuous functions q. It is also important for the matching with recent improvements in the evolution problem to lower the required regularity. This possibility is included in the above theorem, which admits sources* $q \in W^p_{0,\delta+2}$, *and* $a \in W^p_{0,\delta+2}$, $p > \frac{n}{2}$.

We now consider the case when $b \equiv \frac{n-2}{4n}\tau^2 - q_0$ can take negative values. We first prove a calculus lemma.

Lemma 11.9 *Consider the algebraic function of y with* $a > 0, r > 0$, $d \geq 0$

$$f(y) \equiv dy^{\frac{n}{n-2}} - ry^{\frac{n-1}{n-2}} + a \tag{11.21}$$

There are two real numbers y_1 *and* y_2, *such that* $0 < y_1 \leq y_2$ *and*

$$f(y_1) \geq 0, \quad f(y_2) \leq 0 \tag{11.22}$$

if

$$ad^{n-1} < \left[\frac{(n-1)^{n-1}}{n^n}\right] r^n \tag{11.23}$$

Proof Suppose $d > 0$. The function f starts from $a > 0$ for $y = 0$, f decreases when y increases, has a minimum for $y_m = [\frac{(n-1)r}{nd}]^{n-2}$, then f increases up to infinity with y. The equation $f(y) = 0$ has two real roots y_1 and y_2 if the minimum $f(y_m)$ is negative, i.e. the inequality (11.23) is satisfied. Then:

$$f(y) \geq 0, \quad 0 \leq y \leq y_1 \qquad \text{and} \qquad f(y) \leq 0, \quad y_1 \leq y \leq y_2. \tag{11.24}$$

When $d = 0$, $f(y)$ starts then from $a > 0$ and decreases to $-\infty$, $f(y)$ has one root y_1, and satisfies the previous inequalities with $y_2 = +\infty$. □

Theorem 11.10 *Let (M,γ) be a $M^p_{2,\delta}$ manifold, $p > \frac{n}{2}$. Let $a, b, r \in W^p_{0,\delta+2}$ be given on (M,γ), $r \geq 0$, $b \leq 0$ and $\delta > -\frac{n}{p} + \frac{n}{2} - 1$ ($\delta \geq -1$ if $p = 2$). The equation*[63]

$$\triangle_\gamma \varphi - r\varphi + a\varphi^{-\frac{3n-2}{n-2}} - b\varphi^{\frac{n+2}{n-2}} = 0 \tag{11.25}$$

has a solution $\varphi > 0$, with $1 - \varphi \in W^p_{2,\delta}$, if $\delta < n - 2 - \frac{n}{p}$ and if the inequality 11.23 is satisfied on M, with $d = -b$, and if it holds that:

$$\inf_{x\in M} y_1(x) > 0\,, \quad \inf_{x\in M} y_2(x) \geq \max\left\{1, \sup_{x\in M} y_1(x)\right\}, \tag{11.26}$$

where $y_1(x)$ and $y_2(x)$ are the two positive numbers which are roots of the algebraic function[64]

$$f_x(y) \equiv -b(x)y^{\frac{n}{n-2}} - r(x)y^{\frac{n-1}{n-2}} + a(x). \tag{11.27}$$

Proof The equation admits the constant sub- and supersolutions $\varphi_- = \ell > 0$ and $\varphi_+ = m \geq 1$; therefore a solution φ with the given properties exists if for almost every $x \in M$ it holds that

$$f_x(\ell^4) \geq 0, \quad f_x(m^4) \leq 0, \quad \text{with} \quad \ell \leq m. \tag{11.28}$$

The lemma, and the inequalities (11.26), ensure the existence of such numbers ℓ and m, given by:

$$\ell = \min\left\{1, \inf_{x\in M} y_1(x)\right\}, \quad m = \max\{1, \inf_{x\in M} y_2(x)\}. \tag{11.29}$$

□

The second theorem does not rely on the sub–super solution method. It supposes that $r \leq 0$, and hence applies in particular to data sets satisfying the generalized positive Yamabe condition (11.12), since such data sets can be conformally transformed to data satisfying $r = 0$. The theorem has a simple constructive formulation, but it restricts the size of the coefficients a, q, r, and b.

Theorem 11.11 *Let (M,γ) be a $M^p_{2,\delta}$ manifold, $p > \frac{n}{2}$, $\delta > -\frac{n}{p} + \frac{n}{2} - 1$ ($\delta \geq -1$ if $p = 2$). Let $\bar\psi \in W^p_{2,\delta}$, $V(\bar\psi) \in W^p_{0,\delta+2}$, a, q, $b \in W^p_{0,\delta+2}$. Suppose that $b, r \leq 0$. The equation*

$$\triangle_\gamma \varphi - r\varphi + a\varphi^{-\frac{3n-2}{n-2}} + q\varphi^{-\frac{n}{n-2}} - b\varphi^{\frac{n+2}{n-2}} = 0 \tag{11.30}$$

has a solution $\varphi > 0$, with $1 - \varphi \in W^p_{2,\delta}$, if $\delta < n - 2 - \frac{n}{p}$ and if a, b, and r are small enough in $W^p_{0,\delta+2}$ norm.

[63] We have made $q = 0$ to give a simple proof, leaving the more general case for the interested reader to study.

[64] This is a polynomial in the case n=3.

Proof The equation admits the subsolution $\varphi = 1$. We solve it by iteration, starting from $u_0 = 1 - \varphi_0 = 0$. We set:

$$\triangle_\gamma u_1 = -(a+q) + b + r \leq 0 \tag{11.31}$$

and see that u_1 exists, $u_1 \geq 0, u_1 \in W^p_{2,\delta} \subset C^0_\alpha$, and

$$||u_1||_{C^0_\alpha} \leq C_S ||u_1||_{W^p_{2,\delta}} \leq M \quad \text{with} \quad M \equiv C(A+B+R) \tag{11.32}$$

where $C = C_E C_S$ is a number depending only on γ, through the constant C_E of the elliptic estimate and the Sobolev embedding constant C_S and we have set

$$A \equiv ||a+q||_{W^p_{0,\delta+2}}, \quad B \equiv ||b||_{W^p_{0,\delta+2}}, \quad R \equiv ||r||_{W^p_{0,\delta+2}}. \tag{11.33}$$

The inequality (11.32) implies

$$||\varphi_1||_{C^0} \leq 1 + M. \tag{11.34}$$

Recursively, supposing $u_{n-1} \geq 0$ and $||u_{n-1}||_{C^0_\alpha} \leq M$ the equation defining u_n,

$$\triangle_\gamma u_n = -r\varphi_{n-1} + a\varphi_{n-1}^{-\frac{3n-2}{n-2}} + q\varphi_{n-1}^{-\frac{n}{n-2}} - b\varphi_{n-1}^{\frac{n+2}{n-2}}, \tag{11.35}$$

implies $u_n \geq 0$ and

$$||u_n||_{W^p_{2,\delta}} \leq C_E\{A + R(1+M) + B(1+M)^{\frac{n+2}{n-2}}\}. \tag{11.36}$$

hence $||u_n||_{C^0_\alpha} \leq M \equiv C(A+B+R)$ if

$$C\{A + R(1+M) + B(1+M)^{\frac{n+2}{n-2}}\} \leq M \equiv C(A+B+R), \tag{11.37}$$

that is:

$$RM + B[(1+M)^{\frac{n+2}{n-2}} - 1] \leq 0. \tag{11.38}$$

This inequality is satisfied if A, B, R are small enough. The convergence of the series to $u \in W^p_{2,\delta}$ with the $1+u$ solution of (11.30) can be proved by standard methods if A, B, R are small enough. □

Variants of these theorems certainly hold.

12 Solutions of the system of constraints

12.1 *Decoupled system*

The system of the conformally formulated momentum and Hamiltonian constraints decouples, in the asymptotically Euclidean case, when the initial manifold M is maximal and $J_0 = 0$, i.e. when the unscaled matter sources have zero momentum. When the constraints decouple the theorems of the previous sections are sufficient to give existence, non-existence, or uniqueness theorems of the systems of constraints.

Examination of the obtained results gives, for example, the following theorem under a general common hypothesis on *a priori* given data.

Theorem 12.1 ***Hypothesis***. *(M,γ) is a $M^p_{2,\delta}$ manifold; $\tau \equiv 0$; $\bar\psi \in W^p_{2,\delta}$ is a scalar field with potential $V(\bar\psi) \in W^p_{0,\delta+2}$; $\tilde\pi \in W^p_{1,\delta+1}$, $q_0 \in W^p_{0,\delta+2}$ are given scalar fields and $U \in W^p_{1,\delta+1}$ is a symmetric 2-tensor. $J_1 \in W^p_{0,\delta+2}$ is given, $J_0 \equiv 0$. It holds that ($\delta = -1$ is admissible if $p = 2$):*

$$p > \frac{n}{2}, \qquad -1 + \frac{n}{2} - \frac{n}{p} < \delta < -2 + n - \frac{n}{p} \tag{12.1}$$

Conclusion. *The CF system of constraints (10.1) and (11.1) admits a solution X, $\varphi = 1 + u > 0$, with $X, u \in W^p_{2,\delta}$ if the hypotheses of Theorems 11.6, or 11.10, or 11.11 are satisfied.*

Proof We have already proven that under the given hypotheses the constraint (10.1) has a solution, unique, $X \in W^p_{2,\delta}$; therefore $\tilde K \in W^p_{1,\delta+1}$ and $a \in W^p_{0,\delta+2}$ (Sobolev embedding and multiplication; see Appendix I). We know also that $r \in W^p_{0,\delta+2}$. Therefore the coefficients of the Lichnerowicz equation (11.1) satisfy the hypothesis required in the quoted theorems. It has a solution $\varphi > 0$, $\varphi - 1 \in W^p_{2,\delta}$, and the pair X, φ satisfies the CF system of constraints.

This solution is unique in the cases where it is so for the solution of the Lichnerowicz equation. □

12.2 *Coupled system*

We will in this section prove a theorem in the case where the constraints do not decouple, which is in the spirit of a stability theorem. The proof is analogous to the proof given in the case of a compact manifold M, using the implicit function theorem. We consider a solution φ_0, X_0 of the decoupled system and we will discuss the existence of a solution φ and X of the system (10.1), (11.1) of constraints as we perturb τ and J_0 away from zero.

We define as follows a mapping $\mathcal{F}$ from open sets of a pair of Banach spaces into another Banach space:

$$\mathcal{F} : \mathcal{B}_x \times \mathcal{B}_y \to \mathcal{B}_z \text{ by } \ (x; y) \mapsto z, \tag{12.2}$$

with

$$x := (\tau, J_0),\ y := (X, u \equiv \varphi - 1),\ \text{and} \quad z := (\mathcal{H}(x; y),\ \mathcal{M}(x; y)), \tag{12.3}$$

where $\mathcal{H}$ and $\mathcal{M}$ are the left-hand sides of the CF constraints (10.1) and (11.1), and

$$\begin{aligned} &\mathcal{B}_x := W^p_{1,\delta+1} \times W^p_{0,\delta+2}, \qquad \mathcal{B}_y := W^p_{2,\delta} \times W^p_{2,\delta}, \\ &\mathcal{B}_z := W^p_{0,\delta+2} \times W^p_{0,\delta+2}, \quad p > \frac{n}{2}, \ \ \delta > -\frac{n}{p}. \end{aligned} \tag{12.4}$$

The multiplication properties of weighted Sobolev spaces show that $\mathcal{F}$ is a C^1 mapping if φ has no zero. The partial derivative $\mathcal{F}'_y$ is the linear mapping from $\mathcal{B}_y$ into $\mathcal{B}_z$ given by

$$\delta y := (\delta X, \delta u) \mapsto \delta z := (\delta\mathcal{H}, \delta\mathcal{M}). \tag{12.5}$$

The implicit function theorem says that if $(x_0, y_0) \in \mathcal{B}_x \times \mathcal{B}_y$, $\mathcal{F}(x_0, y_0) = 0$ and if the partial derivative $\mathcal{F}'_y$ with respect to y at the point (x_0, y_0) is an isomorphism from $\mathcal{B}_y$ onto $\mathcal{B}_z$, then, there is a neighbourhood Ω of x_0 such that the equation $\mathcal{F}(x, y) = 0$ has one and only one solution y in a neighbourhood of y_0 in $\mathcal{B}_y$ for every $x \in \Omega$.

We can now prove a theorem of existence of solutions with small mean extrinsic curvature, or [and] small unweighted momentum in a neighbourhood of a solution $(\bar{g}_0, K_0)$, such that $tr_{\bar{g}_0} K_0 = 0$, of the system of constraints with data such that $J_0 = 0$. To use the results obtained for the CF constraints, we set $\gamma = \bar{g}_0$, $\varphi_0 = 1$ and decompose K_0 by the formula

$$K_0 \equiv \mathcal{L}_{\bar{g}_0, conf} X_0 + U_0. \tag{12.6}$$

We denote $x_0 = (\tau = 0, J_0 = 0)$ and $y_0 = (X_0, 0)$.

We find that ($\delta b = 0$ because $\tau_0 = 0$ and the unscaled sources are fixed)

$$\delta\mathcal{H} \equiv \Delta_{\bar{g}} \delta u - \alpha_0 \delta u + (\delta a)_0, \tag{12.7}$$

with

$$(\delta a)_0 = \frac{n-2}{2(n-1)} K_0 . \mathcal{L}_{\bar{g}, conf} \delta X \tag{12.8}$$

and

$$\alpha_0 := r_0 + \frac{3n-2}{n-2} a_0 + \frac{n}{n-2} q_1 + \frac{6-n}{n-2} q_2 - \frac{n+2}{n-2} q_0 \, . \tag{12.9}$$

On the other hand

$$(\delta\mathcal{M})_0 \equiv \Delta_{\bar{g}_0, conf} \delta X. \tag{12.10}$$

Theorem 12.2 *Consider a solution $(\bar{g}_0, K_0)$ of the Einstein constraints, $\bar{g}_0 \in M^p_{2,\delta}$, $p > \frac{n}{2}$, $-\frac{n}{p} < \delta < n - 2 - \frac{n}{p}$, $K_0 \equiv \mathcal{L}_{\bar{g}_0, conf} X_0 + U_0$, $X_0 \in W^p_{2,\delta}$ and U_0 a traceless 2-tensor, hence $\tau_0 = 0$, $U_0 \in W^p_{1,\delta+1}$. The sources $\bar{\psi} \in W^p_{2,\delta}$, $V(\bar{\psi}) \in W^p_{0,\delta+2}$, $\tilde{\pi} \in W^p_{1,\delta+1}$, q_0, q_1, q_2, $J_1 \in W^p_{0,\delta+2}$ are assumed to be fixed, except the unscaled current which is zero only for the considered solution.*

Suppose that for some $\tilde{\delta} > -1 + \frac{n}{2} - \frac{n}{p}$ ($\tilde{\delta} = -1$ if $p = 2$) it holds that:

$$\int_M |Df|^2_\gamma + \alpha_0 f^2\} d\mu_\gamma > 0 \quad \textit{for all} \quad f \in W^p_{2,\tilde{\delta}}, \quad f \not\equiv 0 \tag{12.11}$$

with α_0 given above.

Then there exists a neighbourhood Ω of zero in $W^p_{1,\delta+1} \times W^p_{0,\delta+2}$ such that the coupled constraints have one and only one solution (X, φ), with $\varphi > 0$, and $X, u \equiv \varphi - 1 \in W^p_{2,\delta}$ if $(\tau, J_0) \in \Omega$.

Proof The equations

$$\delta\mathcal{M} = f, \quad \text{a given vector in } W^p_{0,\delta+2} \tag{12.12}$$

and

$$\delta\mathcal{H} = h, \quad \text{a given function in } W^p_{0,\delta+2} \tag{12.13}$$

decouple into a well-posed elliptic equation for X, and then a well-posed elliptic equation for δu, under the hypothesis (12.9). The partial derivative of $\mathcal{F}$ with respect to the pair $y := (u, X)$ at the point x_0, y_0 determines therefore an isomorphism from $W^p_{2,\delta} \times W^p_{2,\delta}$ onto $W^p_{0,\delta+2}$, given by

$$(\delta u, \delta X) \mapsto (\delta\mathcal{M}, \delta\mathcal{H}). \tag{12.14}$$

The implicit function theorem proves the result. □

Corollary 12.3 *The inequality (12.9) is always satisfied if $q_0 \leq 0$, hence there are no unscaled matter sources and $V(\bar{\psi}) \leq 0$.*

Proof Since $\varphi_0 = 1$ is a solution of the Lichnerowicz equation for the given data it holds that the value of r for the given solution is

$$r_0 = a_0 + q_1 + q_2 + q_0 \, . \tag{12.15}$$

from which results by an elementary computation[65]

$$\alpha_0 = \frac{4(n-1)}{n-2} a_0 + \frac{2(n-1)}{n-2} q_1 + \frac{4}{n-2} q_2 - \frac{4}{n-2} q_0 \, . \tag{12.16}$$

Therefore $\alpha_0 \geq 0$ if $q_0 \leq 0$. □

13 Gluing solutions of the constraint equations

Section contributed by J. Isenberg and R. Bartnik

Thus far in this chapter, we have primarily discussed methods one can use to generate solutions of the constraints by choosing certain "free data" (e.g. the conformal class of the metric, the mean curvature, and the transverse traceless part of the second fundamental form) and then using the constraint equations to solve for the remaining "determined data" (e.g. the longitudinal part of the second fundamental form and the conformal factor). In a sense, one is constructing solutions of the constraints from scratch.

We now discuss ways of producing new solutions of the constraints from given solutions. Using these "gluing" techniques, one can (a) join a pair of solutions via a connected sum, leaving the old solutions unchanged outside of neighbourhoods of the gluing points; (b) add an arbitrary number of wormholes to a given solution, with no changes outside of a neighbourhood of the wormhole; (c) add an

[65] Remark that the distinction introduced between dimensions $n \geq 6$ and $n < 6$ for the magnetic energy by the conformal scaling disappears here.

arbitrary number of black hole throats to a given solution, with no changes away from neighbourhoods of these throats; and (d) cut away the exterior of a fairly general asymptotically flat solution, and replace the exterior by a solution which, sufficiently far away, is initial data for the Kerr solution. These techniques, in addition to their constructive utility, have proven to be very useful as a tool for proving a number of conjectures of interest.

We first discuss the connected sum (or "IMP") gluing and its generalizations, and we then discuss the exterior (or "CS") gluing.

13.1 *Connected sum gluing*

The idea of the connected sum gluing, as originally developed for the constraint equations by Isenberg, Mazzeo, and Pollack[66] (and hence called "IMP gluing"), is as follows. Say we have two solutions[67] of the constraint equations, (M_1, g_1, K_1) and (M_2, g_2, K_2). Let $p_1 \in M_1$ and $p_2 \in M_2$. Can we find a set of initial data $(M_{(1-2)}, g_{(1-2)}, K_{(1-2)})$ such that (1) M_{1-2} is homotopic to the connected sum[68] $M_1 \# M_2$; (2) $(g_{(1-2)}, K_{(1-2)})$ is a solution of the constraints everywhere on $M_{(1-2)}$; and (3) on that portion of $M_{(1-2)}$ which corresponds to $M_1 \setminus \{\text{ball around } p_1\}$, the data $(g_{(1-2)}, K_{(1-2)})$ is isomorphic to (g_1, K_1), with a corresponding property holding on that portion of $M_{(1-2)}$ which corresponds to $M_2 \setminus \{\text{ball around } p_2\}$? If so, we say that the sets of data admit IMP gluing.

IMP gluing can be carried out for quite general sets of initial data[69]. The sets can be asymptotically Euclidean, asymptotically hyperbolic, specified on a closed manifold, or indeed anything else. The only condition the data sets must satisfy is that, in sufficiently small neighbourhoods of each of the points at which the gluing is to be done, there do not exist non-trivial solutions ξ to the equation $D\Phi^*_{(\gamma,K)}\xi = 0$, where $D\Phi^*_{(\gamma,K)}$ is the linearization operator defined in in equations (2.12–13)) . This "No KIDS" condition signals that the spacetime development of the data does not have any local isometries (see section 2.2).

We note here two features of the proof. First it is constructive, in the sense that it outlines a systematic, step-by-step mathematical procedure for doing the gluing: One conformally blows up the balls surrounding p_1 and p_2 to produce two half cylinders extending from the original initial data sets; one joins the two half cylinders into a bridge, and splices together the data from each side using cutoff functions; one uses the local constant mean curvature to decouple the constraints in the neighborhood of the bridge; one uses tensor projection operators based on linear PDE solutions to find a new conformal K which solves the momentum

[66] Isenberg, J., Mazzeo, R., and Pollack, D. (2002) *Commun. Math. Phys.*, **231**, 529–68, and (2003) *Ann. H. Poincaré*, **4**, 369–83.

[67] In this section, for convenience, we drop the "bar" notation from the solutions of the constraint equations.

[68] The connected sum of these two manifolds is constructed as follows: first we remove a ball from each of the manifolds M_1 and M_2. We then use a cylindrical bridge $S^2 \times I$ (where I is an interval in R^1) to connect the resulting S^2 boundaries on each manifold.

[69] Chruściel, P., Isenberg, I. and Pollack, D. (2005) *Commun. Math. Phys.*, **256**, 29–47.

constraint; one solves the Lichnerowicz equations (3.44) and (3.46) (the argument that this can be done, and that the solution is very close to 1 away from the bridge, relies on the invertibility of the linearized equation, on strong estimates of the constraint equation errors resulting from the splicing stage and on a contraction mapping); one recomposes the data as in Equation (3.6); and finally one does a *non*-conformal data perturbation away from the bridge to return the data there to what it was before the gluing. This procedure can be largely carried out numerically, although we note that it requires us to solve elliptic equations on topologically nontrivial manifolds.

The second feature we note regarding the proof is that it relies primarily on the conformal method, but it also uses a non-conformal deformation of the data at the end to guarantee that the glued data is not just very close to the given data on regions away from the bridge, but is exactly equal to it.

IMP gluing is not the most efficient tool for studying the complete set of solutions of the constraints. It has, however, been found to be very useful for a number of applications, including the following:

1. *Multi-black hole data sets* Given an asymptotically Euclidean solution of the constraints, IMP gluing allows a sequence of flat space initial data sets to be glued to it. The bridges that result from this gluing each contain a minimal surface, and consequently an apparent horizon. With a bit of care[70], one can do this in such a way that indeed the apparent horizons are disjoint, and therefore likely to lead to independent black holes.
2. *Adding a black hole to a cosmological spacetime* Although it is not clear how to define a black hole in a spatially compact solution of Einstein's equations, one can glue an asymptotically Euclidean solution of the constraints to a solution on a compact manifold in such a way that there is an apparent horizon on the bridge. Studying the nature of these solutions of the constraints, and their evolution, could be useful step toward understanding what one might mean by a black hole in a cosmological spacetime.
3. *Adding a* ***wormhole*** *to a given spacetime* While we have discussed IMP gluing as a procedure which builds solutions of the constraints with a bridge connecting two points on different manifolds, it can also be used to build a solution with a bridge connecting a pair of points on the *same* manifold. This allows one to do the following: if one has a globally hyperbolic solution of Einstein's equations, one can choose a Cauchy surface for that solution, choose a pair of points on that Cauchy surface, and glue the solution to itself via a bridge from one of these points to the other. If one now evolves this glued-together initial data into a spacetime, it will likely become singular very quickly because of the collapse of the bridge. Until the singularity develops, however, the solution is essentially as it was before the gluing,

[70] See Chruściel, P. and Delay, E. (2002) *Class. Quant. Grav.*, **19**, L71–L79.

with the addition of an effective wormhole. Hence, this procedure can be used to glue a wormhole onto a generic spacetime solution.

4. *Removing topological obstructions for constraint solutions* We know that every closed three-dimensional manifold M^3 admits a solution of the vacuum constraint equations. To show this, we use the fact that M^3 always admits a metric γ of constant negative scalar curvature. One easily verifies that the data $(\bar{g} = \Gamma, K = \Gamma)$ is a CMC solution. Combining this result with IMP gluing, one can show that for every closed M^3, the manifold $M^3 \setminus \{p\}$ admits both an asymptotically Euclidean and an asymptotically hyperbolic solution of the vacuum constraint equations.
5. *Proving the existence of vacuum solutions on closed manifolds with no CMC (constant mean curvature) Cauchy surface* Based on the work of Bartnik[71], one can show that if one has a set of initial data on the manifold $T^3 \# T^3$ with the metric components even across a central sphere and the components of K odd across that same central sphere, then the spacetime development of that data does not admit a CMC Cauchy surface. Using IMP gluing, one can show that indeed initial data sets of this sort exist.

For certain applications, it might be useful to carry out the gluing operation along blowups of *submanifolds* of the solutions rather than blowups of points. Recent work by Mazzieri (unpublished) shows that so long as the submanifold is at least co-dimension 3 (e.g. a path in four-dimensional solutions of the constraints; nothing bigger than a point in three dimensions), so long as there is an isometry between the embeddings of the submanifold in each set of data, and so long as there are no conformal killing fields present, IMP type gluing can be done. (Actually, the work to date requires that both solutions be on compact manifolds, but it is clear that one will be able to drop this restriction.)

Can IMP gluing be done for solutions of the Einstein constraint equations with non-gravitational fields present? For essentially any Einstein-matter field to which the conformal method can be applied (e.g., Einstein–Maxwell, Einstein–Yang–Mills, Einstein-scaled fluids) gluing can be carried out much as in the vacuum case[72].

13.2 *Exterior (Corvino–Schoen) gluing*

In a certain sense, the form of gluing developed by Corvino and Schoen[73] (and therefore called "CS gluing") is more limited than IMP gluing. While IMP gluing joins a pair of fairly general solutions via connected sum, CS gluing focuses on asymptotically Euclidean solutions and creates new solutions which include a bounded part of the original one in the interior and the complement of a

[71] Bartnik, R. (1988) *Commun. Math. Phys.*, **117**, 615–24.

[72] Isenberg, J., Maxwell, D., and Pollack, D. (2005) *Adv. Theor. Math. Phys.*, **9**, 129–72.

[73] Corvino, J. (2000) *Commun. Math. Phys.*, **214**, 137–89; Corvino, J. and Schoen, R. (2006) *J. Diff. Geom.*, **73**, 85–217. See also Chruściel, P. and Delay, E. (2003) Mémoires de la Société Mathématique de France, **94**, 1–103, and references therein.

bounded portion of a set of Kerr initial data in the exterior. More specifically, let (M^3, g, K) be a smooth, asymptotically Euclidean solution of the constraint equations. If certain asymptotic conditions hold, then for any compact region $\Sigma^3 \subset M^3$ for which $M^3 \setminus \Sigma^3 = \mathbb{R}^3 \setminus B^3$ (where B^3 is a ball in $\mathbb{R}^3$), there is a smooth asymptotically Euclidean solution on M^3 which is identical to the original solution on $\Sigma^3 \subset M^3$, and is identical to Cauchy data for the Kerr solution on $M^3 \setminus \tilde{\Sigma}^3$ for some $\tilde{\Sigma}^3 \subset M^3$. In words, their technique allows us to smoothly glue any interior region of an asymptotically Euclidean solution to an exterior region of a slice of the Kerr solution. For asymptotically Euclidean solutions of the constraints with $K = 0$, this method glues any interior region to an exterior region of a slice of Schwarzschild.

Although CS gluing is limited in application, it is remarkable that it can be carried out: it does involve a transitional solid annular region (with unchosen data) between the chosen interior and the Kerr exterior. However, even allowing for this, it only works because the constraint equations form an underdetermined system, so long as one does not fix the conformal class. In carrying out CS gluing, indeed the conformal class is not fixed.

Before giving some indication of how the result is proven, we note an important application of CS gluing: if we combine CS gluing with certain results of Friedrich, then Chruściel and Delay (2002) show that there is a large class of vacuum spacetime solutions of Einstein's equations which admit complete null infinity regions of the form "scri", as hypothesized by Penrose. The tools developed by Corvino and Schoen have also been used by Chruściel and Delay (2003) to strengthen the IMP gluing results.

The Corvino–Schoen method works by showing that one can solve the constraint equations through a projection using the linearized operator $D\Phi$ and its adjoint $D\Phi^*$ (see the above subsection). To simplify the exposition, we sketch the method in the time-symmetric case, defined by $K = 0$, with the consequences that $\Phi(g, K)$ is replaced by the scalar curvature $R(g)$ and the lapse-shift ξ is replaced by the lapse N. In this case the arguments are essentially the same while the calculations are considerably simpler.

We start with the observation that because DR^* has injective symbol, it satisfies a standard Schauder-type elliptic estimate

$$\|N\|_{H^2(\Omega)} \leq C(\|DR^*N\|_{L^2(\Omega)} + \|N\|_{L^2(\Omega)}), \tag{13.1}$$

on any domain Ω. Since this estimate does not require any control on N at the boundary $\partial\Omega$, it follows that similar estimates hold with weighting functions such as $\rho \in C_c^\infty$ which is positive in Ω and vanishes to high order at $\partial\Omega$:

$$\|N\|^2_{H^2_\rho(\Omega)} := \int_\Omega \rho(N^2 + |\nabla^2 N|^2)\, dv_g \leq C \int_\Omega \rho|DR^*N|^2 + N^2)\, dv_g. \tag{13.2}$$

Here the final term N^2 on the right can be removed if the metric has no Killing vectors. Making this assumption, for all $f \in L^2_\rho(\Omega)$ we can solve $DR(\rho DR^*N) = \rho f$ for $N \in H^4_{\text{loc}}$, which in particular produces a solution to

the linearized constraint equation $DR\,h = \rho f$. An iteration argument may then be used to solve the nonlinear problem $R(g_0+h) = R(g_0)+S$ for any sufficiently small S. This solution $h \in H^2_{\rho^{-1}}(\Omega)$ has the very useful property that it vanishes to high order on $\partial\Omega$. Thus, for example, if $R(g_0)$ is sufficiently small and supported in Ω then there is a perturbation h, also supported in Ω, such that $R(g_0 + h) = 0$.

To use this method to glue a Schwarzschild exterior to an asymptotically flat $R(g) = 0$ metric across an annulus $B_{2R}\backslash B_R$, $R >> 1$, requires one more idea because the flat space kernel $\ker DR^*_\delta = \mathrm{span}(1, x^1, x^2, x^3)$ is non-trivial. This implies that the linearized problem $DR_\delta h = \sigma$ can be solved if and only if σ satisfies the four conditions $\int_\Omega \sigma(1, x^i)d^3x = 0$, and the non-linear problem $R(g_0 + h) = 0$ is similarly obstructed for g_0 close to flat. By choosing the cutoff radius R sufficiently large and rescaling back to $\Omega = B_2\backslash B_1$, we find that we are in this close to flat situation. However, it is possible to solve the projected problem $R(g_0 + h) \in K := \mathrm{span}(1, x^i)$ with uniform estimates on $h \in H^2_{\rho^{-1}}(\Omega)$. This relies on the fact that Schwarzschild exterior metric can be characterized by the mass and centre of mass parameters (m, c^i), defined by

$$g_{\mathrm{Schw}} = \left(1 + \frac{m}{2|\mathbf{x} - \mathbf{c}|}\right)^4 \delta_{ij}. \tag{13.3}$$

Some delicate estimates show that the map $(m, c^i) \mapsto K$ is continuous and has index 1, so there is a choice of parameters (m, c^i) mapping to $0 \in K$, which gives $R(g_0 + h) = 0$ as required. The details of these arguments may be found in the original works[74].

As in the case of IMP gluing, it is very likely that CS gluing works with Maxwell or other non-gravitational fields present. That is, one expects for example that interior regions of asymptotically Euclidean Einstein–Maxwell solutions may generally be glued onto exterior Kerr–Newman initial data sets. This, however, has not yet been proven.

[74] Corvino, J. (2000) *Commun. Math. Phys.*, **214**, 137–89; Corvino, J. and Schoen, R. (2003) gr-qc 0301071; Chruściel, P. and Delay, E. (2003) *Memoires SMF*, **94** and gr-qc 0301073.

VIII

OTHER HYPERBOLIC-ELLIPTIC WELL-POSED SYSTEMS

1 Introduction

The well-posed hyperbolic or hyperbolic-elliptic systems obtained in this chapter lead to the same local existence and geometric uniqueness theorems as the wave gauge choice. However, these different formulations may be important in numerical studies or global existence proofs.

An evolution part of Einstein's equations should exhibit causal propagation, i.e. with the domain of dependence determined by the light cone of the spacetime metric. The non-strict hyperbolic systems considered here lead to local existence and uniqueness only in a Gevrey class[1], but with the dependence domain still determined by the light cone.

2 Leray–Ohya non-hyperbolicity of ${}^{(4)}R_{ij} = 0$

The equations $R_{ij} :=^{(4)} R_{ij} = 0$, called ADM evolution equations, have been largely used in the past for numerical calculations. They are, when N and β are given, a second-order differential system for g_{ij}. The hyperbolicity of a quasilinear system is defined through the linear differential operator obtained by replacing the unknown in the coefficients by given values. In our case for given N, β, and g_{ij} the principal part of this operator acting on a symmetric 2-tensor γ_{ij} is

$$\frac{1}{2}(N^{-2}\partial_{00}^2 - g^{hk}\partial_{hk}^2)\gamma_{ij} + g^{hk}\partial_h\partial_j\gamma_{ik} + g^{hk}\partial_h\partial_i\gamma_{jk} - g^{hk}\partial_i\partial_j\gamma_{hk}.$$

The characteristic matrix at a point of spacetime is the linear operator obtained by replacing the derivation ∂ by a covariant vector ξ. The characteristic determinant is the determinant of this linear operator. We take as independent unknown $\gamma_{12}, \gamma_{23}, \gamma_{31}, \gamma_{11}, \gamma_{22}, \gamma_{33}$ and consider the six equations $R_{ij} = 0$, with the same ordering of indices.

To simplify the writing we compute this matrix in a coframe orthonormal for the given spacetime metric (N, β, g_{ij}). We denote by (t, x, y, z) the components of ξ in such a coframe. The characteristic matrix $\mathcal{M}$ reads then (up to

[1] A Gevrey class is a space of C^∞ functions whose successive derivatives satisfy inequalities weaker than the inequalities satisfied by the derivatives of analytic functions.

multiplication by 2)

$$\begin{pmatrix} t^2-z^2 & zx & z & 0 & 0 & -xy \\ xz & t^2-x^2 & xy & -yz & 0 & 0 \\ yz & xy & t^2-y^2 & 0 & -xz & 0 \\ 2xy & 0 & 2zx & t^2-y^2-z^2 & -x^2 & -x^2 \\ 2xy & 2yz & 0 & -y^2 & t^2-x^2-z^2 & -y^2 \\ 0 & 2yz & 2xz & -z^2 & -z^2 & t^2-x^2-y^2 \end{pmatrix}$$

The characteristic polynomial is the determinant of this matrix. It is found to be

$$Det\mathcal{M} = b^6 a^3, \quad \text{with} \quad b = t, \quad a = t^2 - x^2 - y^2 - z^2.$$

The characteristic cone is the dual of the cone defined in the cotangent plane by setting the characteristic polynomial equal to zero. For our system the characteristic cone splits into the light cone of the given spacetime metric and the normal to its space slice. If the charateristic polynomial admits multiple factors and the system is non-quasidiagonal, it is not hyperbolic in the usual sense. We will prove the following theorem

Theorem 2.1 *When $N > 0$ and β are given and arbitrary, the system $R_{ij} = 0$ is a Leray–Ohya hyperbolic*[2] *system for g_{ij}, in the Gevrey class $\gamma = 2$, as long as g_{ij} is properly Riemannian. If the Cauchy data as well as N and β are in such a Gevrey class, the Cauchy problem has a local in time solution, with domain of dependence determined by the light cone.*

Proof The product by ab^2 of the inverse matrix of the characteristic matrix is computed to be

$$\begin{pmatrix} t^2-x^2-y^2 & -zx & -zy & 0 & 0 & xy \\ -zx & t^2-y^2-z^2 & -xy & zy & 0 & 0 \\ -zy & -xy & t^2-x^2-z^2 & 0 & zx & 0 \\ -2xy & 0 & -2zx & t^2-x^2 & x^2 & x^2 \\ -2xy & -2zy & 0 & y^2 & t^2-y^2 & y^2 \\ 0 & -2zy & -2zx & z^2 & z^2 & t^2-z^2 \end{pmatrix}$$

We see that the elements of the matrix $ab^2\mathcal{M}^{-1}$ are polynomials in x, y, z. The product of this matrix by $\mathcal{M}$ is a diagonal matrix with element in the diagonal ab^2. Consider now the differential operator ${}^{(4)}R_{ij}$ acting on g_{ij}. Multiply it on the left by the differential operator defined by replacing in $ab^2\mathcal{M}^{-1}$ the variables x, y, z by the derivatives $\partial_1, \partial_2, \partial_3$. The resulting operator is quasidiagonal with principal operator $\partial^\lambda \partial_\lambda \partial_0^2$. It is the product of two strictly hyperbolic operators, $\partial^\lambda \partial_\lambda \partial_0$ and ∂_0. The result follows from the general theory of Leray and Ohya[3] (see Appendix IV). □

[2] Called "Hyperbolique non strict" by J. Leray and Y. Ohya.

[3] Leray, J. and Ohya, Y. (1968) *Math. Annalen.*, **162**, 228–36.

3 Wave equation for K

Various hyperbolic systems have been obtained in recent years for the evolution of the dynamical variables (g_{ij}, K_{ij}) by linear combination of R_{ij} with the constraints. Such a hyperbolic system was obtained by Choquet-Bruhat and Ruggeri (1983)[4], in the case of zero shift, and extended in Choquet-Bruhat and York (1995) to the case of an arbitrary shift[5]. It uses a quasidiagonal system of wave equations for K_{ij}, shown to hold modulo a gauge condition for N, called now "densitizing the lapse". It works as follows.

We use the expressions of R_{0i} and R_{ij}, together with

$$\bar{\partial}_0 g_{ik} = -2NK_{ik} \tag{3.1}$$

which imply

$$\bar{\partial}_0 \bar{\Gamma}^h_{ij} = -g^{hk}\{\bar{\nabla}_i(NK_{jk}) + \bar{\nabla}_j(NK_{ik}) - \bar{\nabla}_k(NK_{ij})\} \tag{3.2}$$

to obtain the identity, with $f_{(ij)} = f_{ij} + f_{ji}$, and $H \equiv trace(K) \equiv K^i_i$

$$\begin{aligned}\Omega_{ij} &\equiv \bar{\partial}_0 R_{ij} - \bar{\nabla}_{(i} R_{j)0} \\ &\equiv -\bar{\partial}_0(N^{-1}\bar{\partial}_0 K_{ij}) + \bar{\nabla}_h\bar{\nabla}^h(NK_{ij}) - \bar{\partial}_0(N^{-1}\bar{\nabla}_j\partial_i N) - N\bar{\nabla}_i\partial_j H \\ &\quad + \bar{\partial}_0(HK_{ij} - 2K_{im}K^m_j) - \bar{\nabla}_{(i}(K_{j)h}\partial^h N) - 2N\bar{R}^h_{\ ijm}K^m_h \\ &\quad - N\bar{R}_{m(i}K^m_{j)} + H\bar{\nabla}_j\partial_i N;\end{aligned} \tag{3.3}$$

This identity shows that for a solution of the Einstein equations

$$R_{\alpha\beta} = \rho_{\alpha\beta} \tag{3.4}$$

the extrinsic curvature K satisfies a second-order differential system which is quasidiagonal with principal part the wave operator, except for the terms $-N\bar{\nabla}_i\partial_j H$. The unknown $\bar{g}$ appears at second order, as well as N except for the term $-\hat{\partial}_0(N^{-1}\bar{\nabla}_j\partial_i N)$. One finds that

$$\bar{\partial}_0(N^{-1}\bar{\nabla}_j\partial_i N) + N\bar{\nabla}_j\partial_i H \equiv N^{-1}\bar{\nabla}_j\partial_i(\partial_0 N + N^2 H) + X_{ij} \tag{3.5}$$

where X_{ij} is only of first order in $\bar{g}$ and K.

3.1 *Hyperbolic system*

We densitize the lapse, that is we set[6]

$$N = \alpha(\det \bar{g})^{\frac{1}{2}} \tag{3.6}$$

[4] Choquet-Bruhat, Y. and Ruggeri, T. (1983) *Commun. Math. Phys.*, **89**, 269–75.

[5] A system with a well-defined energy has been obtained by Fritelli, S. and Reula, O. A. (1994) *Commun. Math. Phys.*, **166**, 221–35, by taking just a combination (without derivation) of R_{ij} with the constraints, and densitization of the lapse, in their study of the Newtonian approximation.

[6] Also used with $\alpha = 1$ in the classical paper of Taub on homogeneous cosmologies, reproduced in *Gen. Rel. Grav.*, **36** (2004).

where α is an arbitrary positive tensor density of weight -1 called the "densitized lapse". Then, using the formula for the derivative of a determinant and the identity (3.1), we find

$$\partial_0 N = -N^2 H + \bar{\partial}_0 \alpha (\det \bar{g})^{\frac{1}{2}} \tag{3.7}$$

We see that $\partial_0 N + N^2 H$ is an algebraic function of $\bar{g}$. We have proved the following theorem.

Theorem 3.1 *Let $\alpha > 0$ and β be arbitrary. Set $N = \alpha(\det \bar{g})^{\frac{1}{2}}$, then:*

1. *The equations*

$$\Omega_{ij} = 0 \tag{3.8}$$

   are a quasidiagonal system of wave equations for K.

2. *The equations above together with*

$$\bar{\partial}_0 g_{ij} = -2NK_{ij} \tag{3.9}$$

   are a hyperbolic Leray system[7] *for $\bar{g}$ and K.*

Proof Part 1 has already been proved. To prove part 2 we assign to the equations and unknown the following weights:

$$m(K) = 2, \quad m(\bar{g}) = 2, \quad n(\Omega = 0) = 0, \quad n(\bar{\partial}\bar{g} = 2NK) = 1. \tag{3.10}$$

The principal operator is then a matrix diagonal by blocks. Each block, corresponding to a pair (ij) of indices, is given by

$$\begin{pmatrix} -N^2 \partial_{00}^2 & X \\ 0 & \partial_0 \end{pmatrix} \tag{3.11}$$

with X a second-order operator. The characteristic determinant is $(-N^2 \xi_0^2 + g^{ij}\xi_i \xi_j)\xi_0$, and is a hyperbolic polynomial. □

Remark 3.2 *We have already seen that densitizing the lapse is equivalent to requiring the time to be in wave gauge.*

Known properties of hyperbolic systems on manifolds and the use of the Bianchi identities to show the preservation of constraints under evolution lead to the following theorem, corresponding in the chosen gauge to the local existence theorem proved before in the wave gauge. We formulate it, for a change, in using $H_s^{u,loc}(M)$ spaces.

Theorem 3.3 *Let $(M, \bar{g}_0, K_0)$ be an initial data set satisfying the vacuum constraints, where M is a smooth n-dimensional manifold endowed with a smooth,*

[7] Leray, J. (1953) *Hyperbolic Differential Equations*, Lecture notes, IAS Princeton; see Appendix IV.

Sobolev regular[8]*, Riemannian metric e, and where* $\bar{g}_0 \in H_s^{u.loc}$, $K_0 \in H_{s-1}^{u.loc}$ *are a properly Riemannian metric and a symmetric 2-tensor on* M. *Suppose arbitrarily given on* $M \times I$, I *an interval of* R, *the gauge variables* $\alpha, \beta \in C^0(I, H_s^{u.loc}) \cap C^1(I, H_{s-1}^{u.loc}) \cap C^2(I, H_{s-2}^{u.loc})$. *Then if* $s > \frac{n}{2} + 1$ *there exists an interval* $J \subset I$ *and a Lorentzian metric*

$$g = -N^2 dt^2 + g_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt), \quad N = (\alpha \det \bar{g})^{\frac{1}{2}} \tag{3.12}$$

such that $(M \times J, g)$ *is a solution of the vacuum Einstein equations taking on* M_{t_0}, $t_0 \in J$, *the given initial data* $\bar{g}_0, K_0$.

For a given pair α, β *this solution is unique.*

We already know that (see Chapter 6) that the solution of the Cauchy problem for the vacuum Einstein equations is geometrically unique (i.e. up to isometries) in the class of globally hyperbolic space times.

3.2 *Hyperbolic-elliptic system*

An alternative method to the densitization of the lapse is to consider H as a given function h on space time, i.e. imposing given mean extrinsic curvature on the space slices. The second-order equation for K obtained above reduces again to a quasidiagonal system with principal part the wave operator. This gauge condition was used by Christodoulou and Klainerman (1993)[9], with $h = 0$, in the asymptotically Euclidean case, and in the general case by Choquet-Bruhat and York (1996)[10]. The lapse N is then determined through the equation $R_0^0 = \rho_0^0$ which now reads in the general (non-vacuum) case

$$\bar{\nabla}^i \partial_i N - (|K|^2 - \rho_0^0) N = -\partial_0 h, \quad |K|^2 \equiv K_{ij} K^{ij}$$

This equation is an elliptic equation for N on each time slice when $\bar{g}$, K, and ρ are known.

Note that for energy sources satisfying an energy condition we have $-\rho_0^0 \geq 0$. The coefficient of the undifferentiated term N is therefore non-negative, an important property for the solution of the elliptic equation. The mixed hyperbolic elliptic system that we have constructed will determine the unknowns N and $\bar{g}$ in a neighbourhood of M in $M \times R$ when the shift β is chosen.

As usual for elliptic equations the solution of the equation for N requires a global hypothesis on the manifold M. We state a theorem in the case where M is compact. The case where M is asymptotically Euclidean, used as a local existence theorem in the Christodolou–Klainerman theorem on the non-linear stability of Minkowski spacetime, is treated in Chapter 15.

Theorem 3.4 *Let* $(M, \bar{g}_0, K_0)$ *be an initial data set where* M *is a smooth* n*-dimensional manifold with a smooth Riemannian metric* e. *Suppose that* $\bar{g}_0 \in H_s$,

[8] See Appendix I.

[9] Christodoulou, D. and Klainerman, S. (1993) *Non-Linear Stability of the Minkowski Space*, Princeton University Press.

[10] Choquet-Bruhat, Y. and York, J. W. (1996) (ed. Ferrarese), Pythagora.

$K_0 \in H_{s-1}$ are a properly Riemannian metric and a symmetric 2-tensor on M, and satisfy the vacuum constraints. Then if $s > \frac{n}{2} + 1$ there exists an interval $I \subset R$ and a Lorentzian metric

$$g = -N^2 dt^2 + g_{ij} dx^i dx^j \tag{3.13}$$

such that $(M \times I, g)$ is a solution of the vacuum Einstein equations and takes on M_{t_0}, $t_0 \in I$, the initial data $(M, \bar{g}_0, K_0)$.

Proof The proof is by iteration, solving alternatively the hyperbolic system and the elliptic equation, after choosing a function $h \in C^1(R, H_{s-1})$ such that it takes on M_{t_0} the value trK_0. □

4 Fourth-order non-strict and strict hyperbolic systems for $\bar{g}$

The following fourth-order differential system, satisfied by the time-dependent space metric $\bar{g}$ when the spacetime metric ${}^{(4)}g$ satisfies the vacuum Einstein equations,

$$A_{ij}(\bar{g}) \equiv \partial_{00}^2 R_{ij} - \bar{\partial}_0 \nabla_i R_{j0} - \bar{\partial}_0 \bar{\nabla}_j R_{i0} + \bar{\nabla}_i \bar{\nabla}_j R_{00} = 0$$

has been shown (Anderson *et al.*, 1996) to be Leray–Ohya hyperbolic in the Gevrey class $\gamma = 2$ for arbitrary lapse and shift. It becomes a hyperbolic Leray system by choosing a zero shift and a time wave gauge.

The system $A_{ij} = 0$ plays a role in the study of the propagation of high-frequency gravitational waves (see Chapter 11).

5 First-order hyperbolic systems

Such systems are supposed to be more amenable to numerical computation.

5.1 *FOSH systems*

A first-order system of N equations which reads

$$M^{\alpha IJ}(u)\partial_\alpha u_I + f^J(u) = 0 \tag{5.1}$$

where $u_I, I = 1, \ldots, N$ are a set of unknowns (for instance the components of one or several tensors) is called symmetric if the matrices M^α are symmetric. Such a symmetric system is hyperbolic for the space slices M_t (equivalently we could say "for the time function t") if the matrix $\tilde{M}^t$, the coefficient of $\partial/\partial t$, is positive definite. The energy associated with such systems is straightforward to write in a Cauchy-adapted frame. In this case $\tilde{M}^t = M^0$, while $\tilde{M}^i = M^i - \beta^i M^0$. The equations (5.1) imply

$$u_J M^{\alpha IJ}(u)\partial_\alpha u_I + u_J f^J(u) \equiv \frac{1}{2}\partial_\alpha (M^{\alpha IJ} u_I u_J) + F(u) = 0 \tag{5.2}$$

We express the ∂_α as linear combinations of the usual time and space partial derivatives. The equation above takes the form

$$u_J \tilde{M}^{\alpha IJ}(u) \frac{\partial}{\partial x^\alpha} u_I + u_J f^J(u) \equiv \frac{1}{2} \frac{\partial}{\partial x^\alpha} (\tilde{M}^{\alpha IJ} u_I u_J) + F(u) = 0 \qquad (5.3)$$

Integrating this equation on a strip $M \times [0, T]$ leads to an energy equality and an energy inequality (see Section 4, Appendix IV).

Anderson *et al.* (1995)[11] have written the system (3.1), (3.3), together with the gauge condition (3.6) as a **FOSH** system (first-order symmetric hyperbolic) using also the equation $R_{00} = 0$.

A number of variants have been written since then, and their quality for numerical computation discussed (see in particular the Einstein–Christoffel system of Anderson and York (1999)[12]).

Another definition, in place of symmetry, has been recently used for hyperbolicity of first-order systems: if the number of linearly independent eigenvectors associated to a multiple characteristic is equal to the multiplicity (which is the case for symmetric systems), the system is still called hyperbolic, though it is not always symmetrizable if the dimension of space is greater than 1. The verification of the existence of the relevant number of eigenvectors involves heavy computations. Kidder *et al.* (2001)[13] write a large family of such systems. They show that densitization of the lapse is a necessary condition for the hyperbolicity. They evolve some of them numerically. Proofs of existence theorems involve pseudo differential operators and require further continuity conditions.

6 Bianchi–Einstein equations

The Riemann tensor is the geometric object which intrinsically defines gravitational effects. The introduction of the Bianchi equations satisfied by the Riemann tensor[14] brings, up to now, nothing new for the local existence and uniqueness theorems. It is useful in obtaining geometrical energy estimates (Bel–Robinson energy) leading possibly to global existence theorems. Such estimates have been used by Christodoulou and Klainerman (1992) to prove the global non-linear stability of Minkowski spacetime. They are used by Andersson and Moncrief[15] for global existence of some spacetimes whose space is a compact manifold with negative curvature. The analogue of the Bianchi equations for the Weyl tensor, written in the Newman–Penrose formalism have been used for a long time

[11] Abraham, A., Anderson, A., Choquet-Bruhat, Y., and York, J. (1995) *Phys. Rev. Lett.*, **75**, 3337–81.

[12] Anderson, A. and York, J. (1999) *Phys. Rev. Lett.*, **82**(22), 4384–7.

[13] Kidder, L., Scheel, M., and Teukolsky, S. (2001) *Phys. Rev. D*, **64**(6).

[14] This section is based on the paper by Anderson, A., Choquet-Bruhat, Y. and York, J. (1997) Einstein–Bianchi hyperbolic system for General Relativity, *Topol. Meth. Non-Linear Anal.*, **10**(2), 353–73.

[15] Andersson, L. and Moncrief V. (2004) in *50 Years of the Cauchy Problem* (eds. P. Chruściel and H. Friedrich), Birkhäuser.

by Friedrich[16] to prove global existence of solutions of Einstein equations in a neighbourhood of Minkowski spacetime by an original conformal method.

6.1 *Wave equation for the Riemann tensor*

The Riemann tensor of a pseudo-Riemannian metric, $R_{\alpha\beta,\lambda\mu}$ is antisymmetric in its first pair of indices as well as in its last pair of indices. We call it a symmetric double 2-form because it possesses also the symmetry

$$R_{\alpha\beta,\lambda\mu} \equiv R_{\lambda\mu,\alpha\beta}. \tag{6.1}$$

The Riemann tensor satisfies the Bianchi identities

$$\nabla_\alpha R_{\beta\gamma,\lambda\mu} + \nabla_\gamma R_{\alpha\beta,\lambda\mu} + \nabla_\beta R_{\gamma\alpha,\lambda\mu} \equiv 0. \tag{6.2}$$

These identities imply by contraction

$$\nabla_\alpha R^{\cdots\cdots\alpha}_{\beta\gamma,\ldots,\mu} + \nabla_\gamma R^{\cdots\cdots\alpha}_{\alpha\beta,\ldots,\mu} + \nabla_\beta R^{\cdots\cdots\alpha}_{\gamma\alpha,\ldots,\mu} \equiv 0. \tag{6.3}$$

Using the symmetry (6.1) gives the identities:

$$\nabla_\alpha R^\alpha{}_{\beta,\lambda\mu} + \nabla_\mu R_{\lambda\beta} - \nabla_\lambda R_{\mu\beta} \equiv 0, \tag{6.4}$$

If the Ricci tensor $R_{\alpha\beta}$ satisfies the Einstein equations

$$R_{\alpha\beta} = \rho_{\alpha\beta}, \tag{6.5}$$

then the previous identities imply the equations (Bel[17], Lichnerowicz[18])

$$\nabla_\alpha R^\alpha{}_{\beta,\lambda\mu} = \nabla_\lambda \rho_{\mu\beta} - \nabla_\mu \rho_{\lambda\beta}. \tag{6.6}$$

Equations (6.2) and (6.6) are analogous to the Maxwell equations for the electromagnetic 2-form F:

$$dF = 0, \quad \delta F = J \tag{6.7}$$

where J is the electric current.

Theorem 6.1 *The Riemann tensor of an Einsteinian spacetime of arbitrary dimension satisfies a quasidiagonal semilinear system of wave equations in the spacetime metric.*

Proof One deduces from (6.2) and the Ricci identity an identity of the form:

$$\nabla^\alpha \nabla_\alpha R_{\beta\gamma,\lambda\mu} + \nabla_\gamma \nabla^\alpha R_{\alpha\beta,\lambda\mu} + \nabla_\beta \nabla^\alpha R_{\gamma\alpha,\lambda\mu} + S_{\beta\gamma,\lambda\mu} \equiv 0. \tag{6.8}$$

[16] Friedrich, H. (1981) *Proc. Roy. Soc. Lond.*, **A378**, 401–21 and subsequent papers. See CB-DM2, V 7.
[17] Bel, L. (1958) *C. R. Acad. Sci. Paris.*
[18] Lichnerowicz, A. (1964) *Ann. Sci. IHES*, No. 1.

where $S_{\beta\gamma,\lambda\mu}$ is a homogeneous quadratic form in the Riemann tensor:

$$S_{\beta\gamma,\lambda\mu} \equiv \{R_{\gamma}{}^{\rho} R_{\rho\beta,\lambda\mu} + R^{\alpha}{}_{\gamma,\beta}{}^{\rho} R_{\alpha\rho,\lambda\mu} + [(R^{\alpha}{}_{\gamma,\lambda}{}^{\rho} R_{\alpha\beta,\rho\mu}) - (\lambda \to \mu)]\} - \{\beta \to \gamma\}. \tag{6.9}$$

Using Equations (6.6), when the Ricci tensor satisfies the Einstein equations, gives equations of the form

$$\nabla^{\alpha}\nabla_{\alpha} R_{\beta\gamma,\lambda\mu} + S_{\beta\gamma;\lambda\mu} = J_{\beta\gamma,\lambda\mu}. \tag{6.10}$$

with $J_{\beta\gamma,\lambda\mu}$ depends on the sources $\rho_{\alpha\beta}$ and is zero in vacuum:

$$J_{\beta\gamma,\lambda\mu} \equiv \nabla_{\gamma}(\nabla_{\mu}\rho_{\lambda\beta} - \nabla_{\lambda}\rho_{\mu\beta}) - (\beta \to \gamma). \tag{6.11}$$

□

Moncrief[19] and also Klainerman and Rodnianski[20] have used the elementary solution of the wave equation on the spacetime metric to deduce from the semilinear system satisfied by the Riemann tensor an integral equation satisfied by that tensor. Moncrief uses the Hadamard formulation of the elementary solution, as explained in Friedlander's book, while Klainerman and Rodnianski use the Sobolev construction, as was done, extended to quasilinear equations[21], in the original proof of the local existence theorem for Einstein equations in wave coordinates. It is hoped that these integral formulae will lead to new bounds for the Riemann tensor, like corresponding formulae for the Yang–Mills fields permitted the proof of their L^{∞} boundedness and global existence[22], without restriction on the size of the initial data. Unfortunately there are new difficulties in the case of gravitation, where the spacetime metric is itself an unkown[23]. There is no perspective to a proof of a global existence theorem for generic large data – the singularities theorems (see Chapter 13) say that there is no such theorem for solutions of the Einstein equations, even in vacuum. But it would be interesting to prove the L^2 bounded curvature conjecture made by Klainerman and Rodnianski that a local existence theorem can be proved with only a hypothesis on the local L^2 bound of the Riemann tensor.

6.2 *Case $n = 3$, FOS system*

In a coframe θ^0, θ^i where $g_{0i} = 0$, equations (6.2) with $\{\alpha\beta\gamma\} = \{ijk\}$ and equations (6.6) with $\beta = 0$ do not contain derivatives ∂_0 of the Riemann tensor.

[19] Moncrief, V. (2005) Isaac Newton Institute Preprints.

[20] Klainerman, S. and Rodnianski, I. (2006) arXiv:math.0603009v1, and references therein.

[21] Fourès (Choquet)-Bruhat, Y. (1952) *Acta Mathematica*, **88**, 141–225. See also, for the linear case and estimates, Choquet-Bruhat, Y. (1962) *Annali di Matemetica*, serie 4, **64**, 191–228.

[22] By Eardley and Moncrief on Minkowski spacetime, and by Chruściel and Shatah on a globally hyperbolic manifold; see Chapter 15.

[23] Global existence for large data is not expected from Einstein's equations (see Chapter 13).

We call them "Bianchi constraints". The remaining equations, called from here on "Bianchi equations", read as follows

$$\nabla_0 R_{hk,\lambda\mu} + \nabla_k R_{0h,\lambda\mu} + \nabla_h R_{k0,\lambda\mu} = 0, \tag{6.12}$$

$$\nabla_0 R^0{}_{i,\lambda\mu} + \nabla_h R^h{}_{i,\lambda\mu} = \nabla_\lambda \rho_{\mu i} - \nabla_\mu \rho_{\lambda i} \equiv J_{\lambda\mu i}, \tag{6.13}$$

where the pair $(\lambda\mu)$ is either $(0j)$ or (jl), with $j < l$. There are three of one or the other of these pairs if the space dimension n is equal to 3.

Equations (6.10) and (6.11) are, for each given pair $(\lambda\mu, \lambda < \mu)$, a first-order system for the components $R_{hk,\lambda\mu}$ and $R_{0h,\lambda\mu}$. If we choose at a point of the spacetime an orthonormal frame the principal operator is diagonal by blocks, each block corresponding to a choice of a pair $(\lambda\mu, \lambda < \mu)$ in a symmetric 6×6 matrix which reads:

$$\mathcal{M} = \begin{pmatrix} \partial_0 & 0 & 0 & \partial_2 & -\partial_1 & 0 \\ 0 & \partial_0 & 0 & 0 & \partial_3 & -\partial_2 \\ 0 & 0 & \partial_0 & -\partial_3 & 0 & \partial_1 \\ \partial_2 & 0 & -\partial_3 & \partial_0 & 0 & 0 \\ -\partial_1 & \partial_3 & 0 & 0 & \partial_0 & 0 \\ 0 & -\partial_2 & \partial_1 & 0 & 0 & \partial_0 \end{pmatrix}.$$

We have proved:

Theorem 6.2 *The Bianchi evolution equations take the form, in a Lorentzian orthonormal frame, of a first-order symmetric system.*

The Bianchi equations depend on the choice of frame, as does their hyperbolicity relative to a time function t.

6.3 *Cauchy-adapted frame*

The numerical valued matrix $\mathcal{M}^t$ of coefficients of the operator $\partial/\partial t$ corresponding to the Bianchi equations relative to the Cauchy adapted frame is proportional to the unit matrix, with coefficient N^{-2}, and hence is positive definite and the following theorem holds.

Theorem 6.3 *The Bianchi equations associated with a Cauchy-adapted frame are a FOSH system, with space sections M_t.*

We will give an explicit expression of the full system after introducing two "electric" and two "magnetic" space tensors associated with the double two-form R. They are the gravitational analogues of the electric and magnetic vectors associated with the electromagnetic two-form F. We define the "electric" tensors by

$$E_{ij} \equiv R^0{}_{i,0j}, \tag{6.14}$$

$$D_{ij} \equiv \frac{1}{4} \eta_{ihk} \eta_{jlm} R^{hk,lm}, \tag{6.15}$$

while the "magnetic" tensors are given by

$$H_{ij} \equiv \frac{1}{2} N^{-1} \eta_{ihk} R^{hk}{}_{,0j}, \tag{6.16}$$

$$B_{ji} \equiv \frac{1}{2} N^{-1} \eta_{ihk} R_{0j,}{}^{hk}. \tag{6.17}$$

In these formulae, η_{ijk} is the volume form of the space metric $\bar{g} \equiv g_{ij} dx^i dx^j$.

Lemma 6.4

1. *The electric and magnetic tensors are always such that*

$$E_{ij} = E_{ji}, \quad D_{ij} = D_{ji}, \quad H_{ij} = B_{ji} \tag{6.18}$$

2. *If the Ricci tensor satisfies the vacuum Einstein equations with cosmological constant*

$$R_{\alpha\beta} = \Lambda g_{\alpha\beta} \tag{6.19}$$

then the following additional properties hold

$$H_{ij} = H_{ji} = B_{ij} = B_{ji}, \quad E_{ij} = D_{ij} \tag{6.20}$$

Proof

1. The Riemann tensor is a symmetric double 2-form, the electric and magnetic 2-tensors associated with it by the given relations possess obviously the given symmetries.

2. The **Lanczos identity** for a symmetric double 2-form, with a tilde representing the spacetime double dual, gives

$$\tilde{R}_{\alpha\beta,\lambda\mu} + R_{\alpha\beta,\lambda\mu} \equiv C_{\alpha\lambda} g_{\beta\mu} - C_{\alpha\mu} g_{\beta\lambda} + C_{\beta\mu} g_{\alpha\lambda} - C_{\beta\lambda} g_{\alpha\mu}, \tag{6.21}$$

with

$$C_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{4} R g_{\alpha\beta}. \tag{6.22}$$

It implies that $\tilde{R}_{\alpha\beta,\lambda\mu} + R_{\alpha\beta,\lambda\mu} = 0$ if $C_{\alpha\beta} = 0$, in particular for an Einsteinian vacuum spacetime with possibly a cosmological constant. The relations (6.21) can then be proved by a straightforward calculation that employs the relation $\eta_{0ijk} = N \eta_{ijk}$ between the spacetime and space volume forms. □

In order to extend the treatment to the non-vacuum case and to avoid introducing unphysical characteristics in the solution of the Bianchi equations, we will keep as independent unknowns the four tensors **E**, **D**, **B**, and **H**, which will not be regarded necessarily as symmetric. The symmetries will be imposed eventually on the initial data and shown to be conserved by evolution.

We now express the Bianchi equations in terms of the time-dependent space tensors **E**, **H**, **D**, and **B**. We use the following relations, found by inverting the definitions (6.14–17),

$$R_{0i,0j} = -N^2 E_{ij}, \qquad R_{hk,0j} = N\eta^i_{hk} H_{ij}, \tag{6.23}$$

$$R_{hk,lm} = \eta^i_{hk}\eta^j_{lm} D_{ij}, \quad R_{0j,lm} = N\eta^i_{lm} B_{ji}. \tag{6.24}$$

We will express spacetime covariant derivatives of the Riemann tensor in terms of space covariant derivatives $\bar{\nabla}$ and time derivatives $\bar{\partial}_0$ of **E**, **H**, **D**, **B** by using the connection coefficients in $3+1$ form as given in Chapter 6.

The first Bianchi equation with $[\lambda\mu] = [0j]$ has the form

$$\nabla_0 R_{hk,0j} + \nabla_k R_{0h,0j} - \nabla_h R_{0k,0j} = 0. \tag{6.25}$$

A calculation incorporating previous definitions, then grouping derivatives using $\hat{\partial}_0$ and $\bar{\nabla}_i$, gives to the first pair of Bianchi equations, with $[\lambda\mu] = [0j]$, the following forms:

$$\bar{\partial}_0 E_{ij} - N\eta^{hl}{}_i \bar{\nabla}_h H_{lj} + (L_2)_{ij} = J_{0ji}, \tag{6.26}$$

where **J** is zero in vacuum, and

$$\hat{\partial}_0(\eta^i{}_{hk} H_{ij}) + 2N\bar{\nabla}_{[h} E_{k]j} + (L_1)_{hk,j} = 0, \tag{6.27}$$

with

$$(L_2)_{ij} \equiv -N(\mathrm{tr}K)E_{ij} + NK^k{}_j E_{ik} + 2NK_i{}^k E_{kj} \tag{6.28}$$

$$- (\bar{\nabla}_h N)\eta^{hl}{}_i H_{lj} + NK^k{}_h \eta^{lh}{}_i \eta^m{}_{kj} D_{lm} + (\bar{\nabla}^k N)\eta^l{}_{kj} B_{il}. \tag{6.29}$$

$$(L_1)_{hk,j} \equiv NK^l{}_j \eta^i{}_{hk} H_{il} + 2(\bar{\nabla}_{[h} N)E_{k]j} + 2N\eta^i{}_{lj} K^l{}_{[k} B_{h]i} \tag{6.30}$$

$$- (\bar{\nabla}^l N)\eta^i{}_{hk}\eta^m{}_{lj} D_{im}. \tag{6.31}$$

We see that the non-principal terms L_1 and L_2 are linear in **E**, **D**, **B**, and **H**, with coefficients linear in the geometrical elements $N\mathbf{K}$ and $\bar{\nabla}N$. The characteristic matrix of the principal terms is symmetrizable. The unknowns $E_{i(j)}$ and $H_{i(j)}$, with fixed j and $i = 1, 2, 3$ appear only in the equations with given j. The other unknowns appear in non-principal terms. The characteristic matrix is composed of three blocks around the diagonal, each corresponding to one given j.

The jth block of the characteristic matrix in an orthonormal frame for the space metric $\bar{\mathbf{g}}$, with unknowns listed horizontally and equations listed vertically (j is suppressed) is given by

$$\begin{pmatrix} \xi_0 & 0 & 0 & 0 & N\xi_3 & -N\xi_2 \\ 0 & \xi_0 & 0 & -N\xi_3 & 0 & N\xi_1 \\ 0 & 0 & \xi_0 & N\xi_2 & -N\xi_1 & 0 \\ 0 & -N\xi_3 & N\xi_2 & \xi_0 & 0 & 0 \\ N\xi_3 & 0 & -N\xi_1 & 0 & \xi_0 & 0 \\ -N\xi_2 & N\xi_1 & 0 & 0 & 0 & \xi_0 \end{pmatrix}. \tag{6.32}$$

This matrix is symmetric and its determinant is the characteristic polynomial of the **E**, **H** system. It is given by

$$-N^6(\xi_0\xi^0)(\xi_\alpha\xi^\alpha)^2. \tag{6.33}$$

The characteristic matrix is symmetric in an orthonormal space frame and the coefficient matrix $\mathcal{M}^t$ is positive definite (it is the unit matrix). Therefore, the first-order system is symmetrizable hyperbolic with respect to the space sections M_t. We do not have to compute the symmetrized form explicitly because one can obtain energy estimates directly by using the contravariant associates E^{ij}, $H^{ij}, \ldots$ of the unknowns.

The second pair of Bianchi equations, for D_{ij} and B_{ij}, obtained for $[\lambda\mu] = [lm]$ is analogous. The characteristic matrix for the $[lm]$ equations, with unknowns D_{ij} and B_{ij}, j fixed, with an orthonormal space frame, is the same as the matrix found above.

If the spacetime metric **g** is considered as given, as well as the sources, the Bianchi equations form a linear symmetric hyperbolic system with domain of dependence determined by the light cone of **g**. The coefficients of the terms of order zero are $\bar{\nabla}N$ or $N\mathbf{K}$. The system is homogeneous in vacuum (zero sources).

6.4 *FOSH system for* $\mathbf{u} \equiv (\mathbf{E}, \mathbf{H}, \mathbf{D}, \mathbf{B}, \bar{\mathbf{g}}, \mathbf{K}, \bar{\Gamma})$

The Bianchi equations depend on the metric. Our problem is to find a system for determining the metric from the Riemann tensor (through eventually other auxiliary unknowns), which together with the Bianchi equations, constitute a well-posed system. It is possible to construct a FOSH[24] system linking the metric and the connection to our Bianchi field, if we again densitize the lapse, i.e. set $N = \alpha(\det \bar{\mathbf{g}})^{1/2}$. This system is inspired by an analogous one constructed in conjunction with the Weyl tensor by Friedrich[25].

6.5 *Elliptic-hyperbolic system*

Instead of determining the metric from the curvature through hyperbolic equations one can try to do so by elliptic equations on space slices. Well-posed problems for such equations are essentially global ones and depend on the global geometric properties of the space manifolds.

6.5.1 *Equations for K*

We deduce from the identities (VI.3.18) that K satisfies an equation which resembles a curl equation, namely

$$\bar{\nabla}_j K_{ki} - \bar{\nabla}_k K_{ji} = N^{-1} R_{0i,jk}. \tag{6.34}$$

This equation and the Ricci identity imply that:

$$\bar{\nabla}^j \bar{\nabla}_j K_{ki} - \bar{\nabla}_k \bar{\nabla}^j K_{ji} - \bar{R}_{\ldots k}{}^h K_{hi} - \bar{R}^j_{\ldots ki}{}^h K_{jh} = \bar{\nabla}^j (N^{-1} R_{0i,jk}). \tag{6.35}$$

[24] See Abrahams *et al.* (1996) *loc. cit.*
[25] Friedrich, H. (1996) *Class. Quant. Grav.*, **13**, 1451–9.

As a gauge condition we now suppose that $H := K^h_h$ is a given function h on the spacetime. The use of the momentum constraint gives then in the vacuum case, the equation[26]

$$\bar\nabla^j\bar\nabla_j K_{ki} - \bar R_{\dots k}{}^h K_{hi} - \bar R^j_{\dots ki}{}^h K_{jh} = \bar\nabla_k\partial_i h + \bar\nabla^j(N^{-1}R_{0i,jk}). \tag{6.36}$$

We see that K satisfies (when $\bar g$, N and the Riemann tensor of spacetime are known) a quasidiagonal linear system, elliptic if $\bar g$ is properly Riemannian. This system is formally self-adjoint: its solvability, in some functional space on each manifold M_t, depends on its kernel. The results will rest on properties proved in Appendix II.

Remark 6.5 *When the spacetime metric satisfies the Einstein equation $R_{0j} = 0$, as required in particular in vacuum, the $\bar g$ trace of this equation is, as required for consistency of the argument, the equation*

$$\bar\nabla^j\bar\nabla_j H = \bar\nabla^j\bar\nabla_j h. \tag{6.37}$$

Remark 6.6 *When $n = 3$ the Riemann and Ricci tensor are linked by the identity (vanishing of the Weyl tensor (see Appendix VI))*

$$\bar R^j_{\dots ki}{}^h = \bar g^h_k\bar R^j_i - \bar g^{jh}\bar R_{ik} + \bar g_{ik}\bar R^{jh} - \bar g^j_k\bar R^h_i + \frac{1}{2}(\bar g_{ik}\bar g^{jh} - \bar g^j_i\bar g^h_k)\bar R.$$

Equation (6.36) becomes, with $H = h$

$$\begin{aligned}\bar\nabla^j\bar\nabla_j K_{ki} - \bar R_{\dots k}{}^h K_{hi} - \bar g_{ik}K_{jh}\bar R^{jh} + \frac{1}{2}K_{ik}\bar R \\ = -\frac{1}{2}h\bar g_{ik} + \bar R_{ik}\bar\nabla_k\partial_i h + \bar\nabla^j(N^{-1}R_{0i,jk}).\end{aligned} \tag{6.38}$$

In particular, for a 3-manifold with constant Ricci curvature,

$$\bar R_{ij} = k\bar g_{ij},$$

the equation becomes

$$\bar\nabla^j\bar\nabla_j K_{ik} + \frac{k}{2}K_{ik} = \left(k - \frac{1}{2}\right)h\bar g_{ik} + \bar R_{ik}\bar\nabla_k\partial_i h + \bar\nabla^j(N^{-1}R_{0i,jk}).$$

The only solution of the associated homogeneous equation is $K \equiv 0$ if $k < 0$.

6.5.2 *Equations for the lapse*

The gauge condition $H := Tr_{\bar g}K = h$, a given function, gives, as remarked before, an elliptic equation for the lapse N on each M_t, equivalent to the equation $R_{00} = 0$,

$$\bar\nabla^i\partial_i N - K_{ij}K^{ij}N = -\partial_0 h.$$

[26] Choquet-Bruhat, Y. and York, J. W. (1996) in *Gravitation, Electromagnetism and Geometric Structures* (ed. Ferrarese), Pitagora.

6.5.3 *Equations for* $\bar{g}$

The equation $\bar{\partial}_0 g_{ij} = -2NK_{ij}$ determines as before $\bar{g}$ when N, β, and K are known. However, it does not improve the regularity on M_t of $\bar{g}$ over the regularity of K, and does not lead to a well-posed elliptic-hyperbolic system.

A better result can be sought through the identity VI 3.23 which links the components R_{ijkl} with N, $\bar{g}$ and K, by:

$$\bar{R}_{ij} \equiv \bar{g}^{hk} R_{ih,jk} + HK_{ij} - K_i^h K_{jh} \tag{6.39}$$

Various methods have been devised to determine a Riemannian metric from its Ricci tensor by elliptic equations. The slice harmonic gauge (i.e. harmonic gauge for $\bar{g}$) has been introduced by Andersson and Moncrief[27] to deduce a hyperbolic system from the evolution equations of a $3+1$ decompositions. Indeed formulae analogous to the formulas obtained in Chapter 6 give that:

$$\bar{R}_{ij} \equiv \bar{R}_{ij}^{(e)} + \frac{1}{2}(g_{ih}\bar{\nabla}_j \hat{F}^h + g_{jh}\bar{\nabla}_i \hat{F}^h) \tag{6.40}$$

where $R_{ij}^{(e)}$ is a quasidiagonal system of quasilinear elliptic operator on each M_t of the type

$$\bar{R}_{ij}^{(e)} \equiv -\frac{1}{2}\bar{g}^{hk} D_h D_k g_{ij} + f_{ij}, \tag{6.41}$$

while the vanishing of the vector

$$\hat{F}^h \equiv \bar{g}^{ij}(\Gamma_{ij}^h - \hat{\Gamma}_{ij}^h) = 0 \tag{6.42}$$

expresses gauge harmonicity conditions, that is the fact that the identity mapping from $(M_t, \bar{g}_t)$ into (M, e) is a harmonic map for each t. Such mappings exist globally if (M, e) is a Euclidean space (they are solutions of ordinary Laplace equations), or a compact manifold with negative sectional curvature[28]. The equations

$$\bar{R}_{ij}^{(e)} = \bar{g}^{hk} R_{ih,jk} + HK_{ij} - K_i^h K_{jh} \tag{6.43}$$

together with (6.36) are a quasidiagonal, quasilinear, second-order elliptic system on each slice M_t for $\bar{g}$ and K when N is known.

Since (6.39) are identities, we know that the right-hand side of these equations satisfies equations corresponding to the conservation laws that we saw in previous chapters for the Einstein equations, but now in the properly Riemannian metric $\bar{g}$. The space Bianchi identities applied to $\bar{R}_{ij}$ when $\bar{R}_{ij}^{(e)} = 0$ give for $\hat{F}^h$ the following quasidiagonal homogeneous elliptic system:

$$\bar{\nabla}^j \bar{\nabla}_j \hat{F}^h - \bar{R}_j^h \hat{F}^j = 0. \tag{6.44}$$

[27] Andersson, L. and Moncrief, V. (2004) in *50 Years of the Cauchy Problem* (eds. P. Chruściel and H. Friedrich), Birkhäuser.

[28] Eells, J. and Lemaire, L. (1974) *Bull. Math. Soc.*, **10**, 1–68.

This system has a unique solution $\hat{F}^h = 0$ in a neighbourhood of Euclidean space, or on a compact manifold with negative Ricci curvature. When $\hat{F}^h = 0$ the solution $\bar{g}$ of (6.41) satisfies (6.40). However, we have not proved that a solution $(\bar{g}, K)$ of (6.42), (6.36) satisfies the relation

$$\bar{\partial}_0 g_{ij} \equiv \frac{\partial g_{ij}}{\partial t} - (\bar{\nabla}_i \beta_j + \bar{\nabla}_i \beta_j) = -2NK_{ij}. \tag{6.45}$$

This relation must be compatible with the slice harmonic gauge $\hat{F}^h = 0$ on each M_t, hence compatible with $\frac{\partial}{\partial t}\hat{F}^h = 0$. This will be obtained in the next subsecton through an equation for the shift.

6.5.4 *Equation for the shift*

The definition of $\hat{F}^h$ implies that

$$\frac{\partial}{\partial t}\hat{F}^h \equiv \frac{\partial \bar{g}^{ij}}{\partial t}\bar{\Gamma}^h_{ij} + \bar{g}^{ij}\frac{\partial}{\partial t}\bar{\Gamma}^h_{ij}. \tag{6.46}$$

We have seen in previous chapters, and it is easy to verify again using the expression for $\bar{\Gamma}$ and (6.45), that

$$\frac{\partial}{\partial t}\bar{\Gamma}^h_{ij} = -\{\bar{\nabla}_i(NK^h_j) + \bar{\nabla}_j(NK^h_i) - \bar{\nabla}^h(NK_{ij})\} + \Phi^h_{ij} \tag{6.47}$$

with

$$\Phi^h_{ij} := \frac{1}{2}\{\bar{\nabla}_i(\bar{\nabla}_j\beta^h + \bar{\nabla}^h\beta_j) + \bar{\nabla}_j(\bar{\nabla}_i\beta^h + \bar{\nabla}^h\beta_i) - \bar{\nabla}^h\bar{\nabla}_j\beta_i - \bar{\nabla}^h\bar{\nabla}_i\beta_j\}. \tag{6.48}$$

The use of the Ricci identity gives

$$\Phi^h_{ij} \equiv \frac{1}{2}\{(\bar{\nabla}_j\bar{\nabla}_i + \bar{\nabla}_i\bar{\nabla}_j)\beta^h + \bar{R}_i{}^h{}_{,jk} + \bar{R}_j{}^h{}_{ik}\}\beta^k. \tag{6.49}$$

We see that the equation to satisfy reads:

$$\frac{\partial}{\partial t}\hat{F}^h \equiv 2(-\bar{\nabla}^i\beta^j + NK^{ij})(\bar{\Gamma}^h_{ij} - \hat{\Gamma}^h_{ij}) + \tag{6.50}$$

$$-2\bar{\nabla}_i(NK^{ih}) + \bar{\nabla}^h(NH)\} + \bar{\nabla}^j\bar{\nabla}_j\beta^h + \bar{R}^h{}_k\beta^k = 0 \tag{6.51}$$

Using the momentum constraint $R_0{}^h = 0$ this equation reduces to the linear elliptic equation for β

$$\bar{\nabla}^j\bar{\nabla}_j\beta^h + \bar{R}^h{}_k\beta^k + 2(-\bar{\nabla}^i\beta^j + NK^{ij})(\bar{\Gamma}^h_{ij} - \hat{\Gamma}^h_{ij}) - 2K^{ih}\bar{\nabla}_i N - H\bar{\nabla}^h N - 2N\bar{\nabla}H = 0. \tag{6.52}$$

This equation has good properties on compact manifolds $(M, \bar{g})$ with negative Ricci curvature, or on asymptotically Euclidean manifolds with small or non-positive Ricci curvature.

7 Bel–Robinson tensor and energy

An energy functional on each space slice corresponds to the FOSH system of the Bianchi equations that we have written in a Cauchy-adapted frame. We will give explicitly this energy, and the inequality it satisfies, in the next section. In the present section we give a spacetime 4-tensor introduced by Bel[29], which can be considered as a "higher order stress energy tensor" from which the Bel–Robinson energy can be deduced.

7.1 *The Bel tensor*

The **Bel tensor** associated with a double 2-form **A** on a four-dimensional pseudo-Riemannian manifold has been defined by Bel through the use of the left and right adjoints, and the biadjoint of **A**. The left and right adjoints are double 2-forms given in an arbitrary frame by:

$$({}^*A)_{\alpha\beta,\lambda\mu} \equiv \frac{1}{2}\eta_{\alpha\beta\rho\sigma}A^{\rho\sigma}_{\ldots\ldots,\lambda\mu} \tag{7.1}$$

$$(A^*)_{\alpha\beta,\lambda\mu} \equiv \frac{1}{2}\eta_{\rho\sigma\rho\sigma}A^{\ldots\ldots\rho\sigma}_{\alpha\beta,} \tag{7.2}$$

with $\eta_{\alpha\beta\lambda\mu}$ the volume form of the spacetime metric. If the double 2- form is symmetric its left and right adjoints are equal. The biadjoint is the double 2-form:

$$({}^*A^*)_{\alpha\beta,\lambda\mu} \equiv \frac{1}{4}\eta_{\alpha\beta\rho\sigma}\eta_{\gamma\delta\lambda\mu}A^{\rho\sigma,\gamma\delta} \tag{7.3}$$

It is symmetric if **A** is symmetric.

Lemma 7.1 *If* **A** *is a symmetric double 2-form it is linked to its biadjoint by the* ***Lanczos identity***

$$(A +^* A^*)_{\alpha\beta,\lambda\mu} \equiv B_{\alpha\lambda}g_{\beta\mu} + B_{\beta\mu}g_{\alpha\lambda} - B_{\alpha\mu}g_{\beta\lambda} - B_{\beta\lambda}g_{\alpha\mu} \tag{7.4}$$

with

$$B_{\alpha\lambda} \equiv A_{\alpha\lambda} - \frac{1}{4}g_{\alpha\lambda}A \tag{7.5}$$

where $A_{\alpha\lambda}$ *and* A *are respectively the 2-tensor and scalar obtained by contraction of* **A** *with the metric.*

Proof Straigtforward computation, easiest to do in an orthonormal frame. □

Corollary 7.2 *If* **A** *is the Riemann tensor of a vacuum Einsteinian spacetime, with possibly a cosmological constant, it is equal to its biadjoint.*

[29] Bel, L. (1958) *loc. cit.*

Definition 7.3 *The* ***Bel tensor***[30] $Q(A)$ *of a double 2-form* A *is the fourth-order tensor given by*

$$Q^{\alpha\beta\lambda\mu}(A) \equiv \frac{1}{4}\{A^{\alpha\rho,\lambda\sigma}A^{\beta\ldots,\mu}_{\ldots\rho,\ldots\sigma} + (^*A)^{\alpha\rho,\lambda\sigma}(^*A)^{\beta\ldots,\mu}_{\ldots\rho,\ldots\sigma} + (A^*)^{\alpha\rho,\lambda\sigma}(A^*)^{\beta\ldots,\mu}_{\ldots\rho,\ldots\sigma} + (^*A^*)^{\alpha\rho,\lambda\sigma}(^*A^*)^{\beta\ldots,\mu}_{\ldots\rho,\ldots\sigma} \tag{7.6}$$

It is symmetric in its two first and two last indices, and by commutation of these pairs of indices if A *is a symmetric double 2-form.*

Theorem 7.4 *If* A *is a symmetric double 2-form with vanishing covariant differential, i.e. such that (the sum is over circular permutation of indices)*

$$\sum_{\alpha\beta\gamma} \nabla_\alpha A_{\beta\gamma,\lambda\mu} \equiv 0 \tag{7.7}$$

then the Bel tensor of A *satisfies an identity of the form*

$$\nabla_\alpha Q^{\alpha\beta\lambda\mu}(A) \equiv L^{\beta\lambda\mu}(A) \tag{7.8}$$

where $L^{\beta\lambda\mu}$ *is linear and homogeneous in the covariant derivatives of the 2-tensor* $A_{\alpha\beta}$*.*

Proof Straightforward calculation using previous relations. □

7.2 *The Bel–Robinson tensor and energy*

The Riemann tensor of a pseudo-Riemannian metric is a symmetric double 2-form satisfying the identities (7.7) (Bianchi identities). Its Bel tensor, called in this case the **Bel–Robinson tensor**, simply denoted Q, satisfies the relations (7.8); hence the following theorem.

Theorem 7.5 *The Bel–Robinson tensor of a vacuum Einsteinian spacetime (possibly with cosmological constant) is conservative, namely it satisfies the equation*

$$\nabla_\alpha Q^{\alpha\beta\lambda\mu} = 0. \tag{7.9}$$

It can be proved that the contraction of the Bel–Robinson tensor with timelike vectors is positive, giving thus the definition of a positive energy density for the Riemann tensor, which leads through the conservation equations (7.9) to energy inequalities.

[30] We choose the notation Q to be in agreement with Christodoulou–Klainerman notations (see Chapter 15).

8 Bel–Robinson energy in a strip

We return to the Bianchi equations written in a Cauchy adapted frame.

Multiply (6.27) by $\frac{1}{2}\eta_l{}^{hk}H^{lj}$ and recall that $\eta_l{}^{hk}\eta^i{}_{hk} = 2\delta^i{}_l$, $\eta_{lrk}\eta^{ihk} = \delta^i{}_l\delta^h{}_r - \delta^i{}_r\delta^h{}_l$, and $\bar\partial_0 g^{ij} = 2NK^{ij}$. Then we find that

$$\frac{1}{2}\eta_l{}^{hk}H^{lj}\bar\partial_0(\eta^i{}_{hk}H_{ij}) = \frac{1}{2}\bar\partial_0(H_{ij}H^{ij}) - M_1, \tag{8.1}$$

$$\begin{aligned} M_1 &\equiv \frac{1}{4}\eta^l{}_{rs}H_{lm}\eta^i{}_{hk}H_{ij}\bar\partial_0(g^{hr}g^{ks}g^{jm}) \\ &= N(K^h_hH^{ij} - K^i{}_lH^{lj} + K_l{}^jH^{il})H_{ij}. \end{aligned} \tag{8.2}$$

Likewise, multiply (6.26) by E^{ij} to obtain

$$E^{ij}\bar\partial_0E_{ij} = \frac{1}{2}\bar\partial_0(E_{ij}E^{ij}) - M_2, \tag{8.3}$$

$$M_2 \equiv N(K^i{}_lE^{lj} + K_l{}^jE^{il})E_{ij}. \tag{8.4}$$

Multiplication by appropriate factors[31] of the second pair of Bianchi equations leads to analogous results. The sum of the expressions so obtained from the four Bianchi equations gives an expression where the spatial derivatives add to form an exact spatial divergence, just as for all symmetric systems. Indeed, we obtain

$$\begin{aligned} &\frac{1}{2}\partial_0(|\mathbf{E}|^2 + |\mathbf{H}|^2 + |\mathbf{D}|^2 + |\mathbf{B}|^2) + \bar\nabla_h(NE^{ij}\eta^{lh}{}_iH_{lj}) \\ &\quad - \bar\nabla_h(NB^{ij}\eta^{lh}{}_iD_{lj}) = Q(\mathbf{E},\mathbf{H},\mathbf{D},\mathbf{B}) + \mathcal{S}, \end{aligned} \tag{8.5}$$

where we have denoted by $|\cdot|$ the pointwise $\bar{\mathbf{g}}$ norm of a space tensor, and where Q is a quadratic form with coefficients $\bar\nabla N$ and $N\mathbf{K}$. The source term $\mathcal{S}$ is

$$\mathcal{S} \equiv J_{0ij}E^{ij} - \frac{1}{2}NJ_{lmi}\eta_h{}^{lm}B^{ih}, \tag{8.6}$$

and is zero in vacuum.

We define the ***Bel–Robinson energy*** at time t of the field $(\mathbf{E},\mathbf{H},\mathbf{D},\mathbf{B})$, called a "Bianchi field" when it satisfies the Bianchi equations, to be the integral

$$\mathcal{B}(t) \equiv \frac{1}{2}\int_{M_t}(|\mathbf{E}|^2 + |\mathbf{H}|^2 + |\mathbf{D}|^2 + |\mathbf{B}|^2)\mu_{\bar{\mathbf{g}}_t}. \tag{8.7}$$

We will prove the following.

Theorem 8.1 *Suppose that* $\mathbf{g}$ *is* $\mathcal{C}^1$ *on* $M\times[0,T]$ *and that the* $\bar{\mathbf{g}}_t$ *norms of* $\bar\nabla N$ *and* $N\mathbf{K}$ *are uniformly bounded on* M_t, $t\in[0,T]$. *Denote by* $\pi(t)$ *the supremum*

$$\pi(t) = \mathrm{Sup}_{M_t}(|\bar\nabla N| + |N\mathbf{K}|). \tag{8.8}$$

[31] Anderson, A., Choquet-Bruhat, Y., and York, J. W. (1997) *loc. cit.*

Suppose the matter source $\mathbf{J} \in L^1([0,T], L^2(M_t))$*; then the Bel energy of a* $\mathcal{C}^1$ *Bianchi field with compact support in space satisfies for* $0 \leq t \leq T$ *the following inequality*

$$\mathcal{B}(t)^{\frac{1}{2}} \leq \left(\mathcal{B}(0)^{\frac{1}{2}} + \frac{C}{2} \int_0^t \|J\|_{L^2(M_\tau)} d\tau \right) \exp\left(C \int_0^t \pi(\tau) d\tau \right), \tag{8.9}$$

where C *is a given positive number.*

Proof We integrate the identity (8.5) on the strip $M \times [0,t]$ with respect to the volume element $\mu_{\bar{\mathbf{g}}_\tau} d\tau$. If the Bianchi field has support compact in space the integral of the space divergence term vanishes. The integration of $\partial_0 f$ on a strip of spacetime with respect to the volume form $d\tau\, \mu_{\bar{\mathbf{g}}_\tau}$ goes as follows, for an arbitrary function f,

$$\int_0^t \int_{M_\tau} \partial_0 f \, \mu_{\bar{\mathbf{g}}_\tau} d\tau \equiv \int_0^t \int_{M_\tau} (\partial_t - \beta^i \partial_i) f \, \mu_{\bar{\mathbf{g}}_\tau} \, d\tau.$$

One has

$$\int_{M_t} \partial_0 f \, \mu_{\bar{\mathbf{g}}_t} = \int_{M_t} \{ \partial_t (f \mu_{\bar{g}_t}) - f \partial_t \, \mu_{\bar{\mathbf{g}}_t} \} - \{ \bar{\nabla}_i (\beta^i f) + f \bar{\nabla}_i \beta^i \} \, \mu_{\bar{\mathbf{g}}_t}. \tag{8.10}$$

We find, using the expression for the derivative of a determinant and the relation between $\bar{g}$ and K, that

$$\partial_t \mu_{\bar{g}_t} = (-NH + \bar{\nabla}_i \beta^i) \mu_{\bar{g}_t}, \qquad H := trK. \tag{8.11}$$

Therefore, if f has compact support in space

$$\int_{M_t} \partial_0 f \, \mu_{\bar{\mathbf{g}}_t} = \frac{d}{dt} \int_{M_t} f \mu_{\bar{g}_t} + \int_{M_t} f N H \mu_{\bar{\mathbf{g}}_t}. \tag{8.12}$$

The integration on a strip leads therefore to the equality

$$\mathcal{B}(t) = \mathcal{B}(0) + \int_0^t \int_{M_\tau} (\tilde{Q} + \mathcal{S}) \, \mu_{\bar{\mathbf{g}}_\tau} \, d\tau, \tag{8.13}$$

with

$$\tilde{Q} = Q - \frac{1}{2} N H (|\mathbf{E}|^2 + |\mathbf{H}|^2 + |\mathbf{D}|^2 + |\mathbf{B}|^2).$$

The expression for Q is deduced from the values of M_1 and M_2 given by (8.2) and (8.4). The equality (8.13) leads therefore to an inequality of the following type, with C some number:

$$\mathcal{B}(t) \leq \mathcal{B}(0) + C \left\{ \int_0^t \mathrm{Sup}_{M_\tau} (|\bar{\nabla} N| + |N\mathbf{K}|) \mathcal{B}(\tau) \, + \|J\|_{L^2(M_\tau)} \mathcal{B}(t)^{\frac{1}{2}} \right\} d\tau. \tag{8.14}$$

This inequality and the resolution of the corresponding equality imply the result, using the Gromwall lemma. □

Remark 8.2 *The quantities* $-\bar{\nabla}_k N$ *and* $-2NK_{ij}$ *are, respectively, the* $(0k)$ *and* (ij) *components of the Lie derivative* $\mathcal{L}_n \mathbf{g}$ *of the spacetime metric* $\mathbf{g}$ *with respect to the unit normal* n *to* M_t*, the* (00) *component of* $\mathcal{L}_n \mathbf{g}$ *is identically zero. The Bel–Robinson energy, therefore, is conserved if this Lie derivative is zero, that is if* $\mathbf{g}$ *is invariant under the flow of* n*.*

The estimate of the Bel–Robinson energy is only an intermediate step in global existence proofs, since it depends on the metric, which itself depends on curvature. We will return to the utilization of the Bel–Robinson energy in Chapters 15 and 16 on global existence problems.

IX

RELATIVISTIC FLUIDS

1 Introduction

A **fluid source** in a domain of a spacetime (V, g) is such that there exists in this domain a unit timelike vector field u, satisfying $g(u,u) \equiv g_{\alpha\beta}u^\alpha u^\beta = -1$, whose trajectories are the flow lines of matter. A moving Lorentzian orthonormal frame is called a **proper frame** if its timelike vector is u.

One may also consider null fluids, with flow lines trajectories of a null vector field u, i.e. such that $u^\alpha u_\alpha = 0$.

Suppose that for a given type of matter there exists in Special Relativity a symmetric stress energy momentum tensor T which is local. By this we mean that T depends only on spacetime functions or fields including u, and possibly their derivatives. One then argues from the equivalence principle in order to carry over to General Relativity the expression obtained for the stress energy tensor in Special Relativity and arbitrary coordinates.

The Einstein equations are always

$$S_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = T_{\alpha\beta} \tag{1.1}$$

The conservation equations,

$$\nabla_\alpha T^{\alpha\beta} = 0, \tag{1.2}$$

which in Special Relativity result from the physical laws of conservation of energy and momentum, are now a consequence of the Bianchi identities. This property was a motivation for Einstein in the choice of his equations.

The conservation equations (1.2) must often be completed by other equations satisfied by the physical quantities appearing in T.

In this chapter we will give general properties of perfect fluids. When appropriate definitions are given, some of these properties generalize well-known properties of non-relativistic perfect fluids. However, the equivalence of mass and energy in Relativity introduces some fundamental differences. The limitation by the speed of light of the speed of any macroscopically transmitted signal also leads to new considerations.

We will touch only briefly on the case of dissipative fluids.

2 Case of dust

In the case of dust, also called pure matter or incoherent matter, there are neither momentum nor stresses in a proper frame and the stress energy tensor reads, as in Special Relativity,

$$T_{\alpha\beta} = r u_\alpha u_\beta, \qquad u^\alpha u_\alpha = -1, \tag{2.1}$$

where r is the proper mass density and u the unit flow velocity. The scalar r is often called the particle number density, a name particularly appropriate in the case of null fluids.

At the usual laboratory, and sometimes astronomical, scale, this dust stress energy tensor is, in some contexts, a good approximation of a general fluid stress energy tensor, since the proper mass energy is of order c^2 (c the velocity of light), with respect to other forms of energy.

2.1 *Equations*

The conservation laws for the dust stress energy tensor read:

$$\nabla_\alpha T^{\alpha\beta} \equiv u_\beta \nabla_\alpha (r u^\alpha) + r u^\alpha \nabla_\alpha u^\beta = 0. \tag{2.2}$$

They give, using the property $u^\beta u_\beta = -1$, hence $u_\beta \nabla_\alpha u^\beta = 0$, the **continuity equation** (conservation of matter)

$$\nabla_\alpha (r u^\alpha) = 0 \tag{2.3}$$

and the **geodesic motion** of the particles

$$u^\alpha \nabla_\alpha u^\beta = 0. \tag{2.4}$$

Similar equations are obtained for a **null dust** model where $u^\alpha u_\alpha = 0$.

The geometric initial data for the spacetime metric g on an initial manifold M are a Riemannian metric $\bar{g}$ and a symmetric 2-tensor K. The initial data for a dust source are a scalar function $\bar{r}$ on M and a tangent vector field $\bar{v}$ to M. A set (V, g, r, u) solution of the coupled Einstein-dust equations is an Einsteinian development of the initial data set $(M, \bar{g}, K, \bar{r}, \bar{v})$ if $\bar{g}$ and K are respectively the induced metric and second fundamental form of M as embedded submanifold in (V, g), while $\bar{r}$ is the function induced by r on M and $\bar{v}$ is the value on M of the dust velocity with respect to the proper frame of an observer with timelike vector orthogonal[1] to M in (V, g). In local coordinates such that the values on M of the shift and the lapse of the development are respectively $\bar{\beta} = 0$ and $\bar{N} = 1$, it holds that $\bar{v}^i = (\bar{u}^0)^{-1} \bar{u}^i$, where $\bar{u}^\alpha$ are the components of u in the considered coordinate system at points of M.

[1] Two such timelike vectors corresponding to two developments isometric under a diffeomorphism f are mapped on to each other by f.

We use the following definition[2] of the spaces $E_s(T)$

$$E_s(T) := C^0([0,T], H_s) \cap C^1([0,T], H_{s-1}) \cap C^2([0,T], H_{s-2}).$$

We denote by $L_s(T)$ the space of Lorentzian metrics which are continuous and bounded on $M \times [0,T]$ with $\partial g \in E_{s-1}(T)$. We denote by $U_s(T)$ the space of unit vectors in the metric g which are continuous and bounded on $M \times [0,T]$ with $\partial u \in E_{s-1}(T)$. We will prove the following existence and uniqueness theorem[3].

Theorem 2.1 *An initial data set $(M, \bar{g}, K, \bar{r}, \bar{v})$ for Einstein equations with dust source admits a globally hyperbolic causal Einsteinian development (V, g, r, u) if $\bar{g} \in M_s$, $K \in H_{s-1}$, $\bar{v} \in H_{s-1}$, $|\bar{v}|_{\bar{g}} < 1$, $\bar{r} \in H_{s-2}$, $s > \frac{n}{2} + 2$, and the constraints are satisfied. There is a number $T > 0$ such that $g \in L_s(T)$, $u \in U_{s-1}(T)$, $r \in E_{s-2}(T)$. The development is unique up to isometries in the class of maximal globally hyperbolic developments if $s > \frac{n}{2}+3$, conjecture $> \frac{n}{2}+2$.*

Proof The Einstein equations with dust source read in wave coordinates:

$$R^{(h)}_{\alpha\beta} \equiv -\frac{1}{2} g^{\lambda\mu} \partial^2_{\lambda\mu} g_{\alpha\beta} + H_{\alpha\beta}(g, \partial g) = \rho_{\alpha\beta} \equiv r \left(u_\alpha u_\beta + \frac{1}{n-1} g_{\alpha\beta} \right). \qquad (2.5)$$

They are to be coupled with Equations (2.4) and (2.3). The initial values for the components of g and u in coordinates are deduced from the geometric data $(\bar{g}, \bar{v}, K)$ by choosing wave coordinates such that the values on M of the lapse and the shift of the metric are respectively $\bar{N} = 1$ and $\bar{\beta} = 0$ (compare with the vacuum case, Chapter 6).

We differentiate Equations (2.5) in the direction of u and use Equations (2.4) and (2.3) to obtain equations which are of third order in g, but still do not contain derivatives of r:

$$u^\gamma \nabla_\gamma R^{(h)}_{\alpha\beta} = -r \nabla_\gamma u^\gamma \left(u_\alpha u_\beta + \frac{1}{n-1} g_{\alpha\beta} \right). \qquad (2.6)$$

The initial values for the second derivatives of g are determined by Equations (2.5).

We give to the equations the Leray–Volevic indices

$$m(2.5) = 0, \quad m(2.4) = 1, \quad m(2.3) = 0 \qquad (2.7)$$

and to the unknowns g, u, r the indices

$$\ell(g) = 3, \quad \ell(u) = 2, \quad \ell(r) = 1 \qquad (2.8)$$

[2] H_s denotes a usual Sobolev (Hilbert) space of functions or tensors on M which are square integrable together with their derivatives of order up to s, in a given smooth, Sobolev regular (for instance complete) Riemannian metric on the manifold M. The notation $M_s$stands for continuous and bounded Riemannian metrics with derivatives in H_{s-1}. On non-compact manifolds, variants of these spaces can be used (see Appendix I).

[3] Remark that this theorem makes no hypothesis on the sign of $\bar{r}$. In particular it can vanish in some domains.

The matrix of principal parts, of order the various $\ell - m$, is then diagonal, given by

$$\begin{pmatrix} -\frac{1}{2}g^{\lambda\mu}u^{\gamma}\partial^{3}_{\gamma\lambda\mu} & 0 & 0 \\ 0 & u^{\alpha}\partial_{\alpha} & 0 \\ 0 & 0 & u^{\alpha}\partial_{\alpha} \end{pmatrix} \tag{2.9}$$

The system is quasidiagonal with characteristic cotangent cone

$$g^{\lambda\mu}\xi_{\lambda}\xi_{\mu}u^{\alpha}\xi_{\alpha} = 0, \tag{2.10}$$

the union of the light cone and a spacelike hyperplane exterior to it if u is a timelike vector. The system is then Leray-hyperbolic and causal. The coefficients of the principal terms are in $C^{1,1}$ under the given hypothesis. The Leray–Dionne theory gives, after some work[4], the existence of a number $T > 0$ such that the considered Cauchy problem has a unique solution $g \in L_s(T)$, $u \in U_{s-1}(T)$, $r \in E_{s-2}(T)$.

The proofs of geometric local existence, and global uniqueness, independent of coordinates, follow the same lines as in the vacuum case. □

2.2 *Motion of isolated bodies (Choquet–Bruhat and Friedrich, 2006)*

The problem of how to use the general equations of fluid evolution to determine the motion of isolated bodies is a longstanding problem which has received only very partial answers. In the pure matter case, we will see that the situation is more satisfactory.

Consider initial data $\bar{g}, K, \bar{v}, \bar{r}$ on M satisfying the hypothesis of the existence and uniqueness in Theorem 2.1. Suppose the support of the initial density $\bar{r}$ is the closure $\bar{\omega}$ of some open subset $\omega \subset M$, which can have several disconnected components, with spaces occupied by material bodies at the initial time.[5] We have just proved that the Cauchy problem with data $\bar{g}, K, \bar{v}, \bar{r}$ is geometrically well posed; the motion of our isolated bodies is therefore determined by these Cauchy data. We want now to prove that this motion, i.e. the solution we have found, does not depend on the "unphysical" data, that is the data of $\bar{v}$ on the subset where $\bar{r}$ vanishes. To prove this property we will use the facts that, on the one hand, the support of r is contained in the geodesic tube Ω issued from the support of $\bar{r}$, and that, on the other hand, the right-hand side of (2.5) vanishes outside of the support of r.

Theorem 2.2 *The values of g, r, on V and u in Ω, the geodesic tube of the metric g based on the support $\bar{\omega}$ of $\bar{r}$, of the solution of Theorem 2.1 do not depend on the initial value $\bar{v}$ in the domain $M\backslash\bar{\omega}$.*

[4] The regularity of data required by the general Leray theory of hyperbolic systems was improved by Dionne only in the case of one equation, though of arbitrary order. The technical details for extension to systems would be out of place in this book.

[5] Remark that the density tends continuously to zero at the boundary since Theorem 2.1 supposes $\bar{r} \in H_2$.

Proof Let g, r, u be the solution with initial values $(\bar{g}, K, \bar{r}, \bar{v})$, *Support* $(\bar{r}) \subset \bar{\omega}$ obtained in Theorem 2.1. Equation (2.3) shows that the support of r is contained in the geodesic tube Ω based on the support $\bar{\omega}$ of $\bar{r}$, generated by geodesics tangent to $\bar{u}$ issued from points in $\bar{\omega}$. Consider another initial value $\bar{v}_1$ for u, such that $\bar{v}_1 = \bar{v}$ on $\bar{\omega}$. The geodesic tube Ω_1 corresponding to $\bar{v}_1$ coincides with Ω; there $u = u_1$, and the dust equations are satisfied in Ω. Since $r = 0$ outside of Ω, the dust stress energy tensors $ru \otimes u$ and $ru_1 \otimes u_1$ coincide on the whole of M. The triple (g, r, u_1) satisfies the Einstein dust system on V, with $u_1 = u$ on $V \backslash \Omega$. The geometrical uniqueness in Theorem 2.1 completes the proof. □

3 Charged dust

The stress energy tensor of charged pure matter (dust) is the sum of the stress energy tensor of the matter and the **Maxwell tensor** of the electromagnetic field F:

$$T_{\alpha\beta} = ru_\alpha u_\beta + \tau_{\alpha\beta}, \tag{3.1}$$

with

$$\tau_{\alpha\beta} = F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}. \tag{3.2}$$

3.1 *Equations*

The Einstein equations read

$$R_{\alpha\beta} = \Phi_{\alpha\beta} + r\left(u_\alpha u_\beta + \frac{1}{n-1} g_{\alpha\beta}\right) \tag{3.3}$$

where we have set

$$\Phi_{\alpha\beta} := F_\alpha{}^\lambda F_{\beta\lambda} + c_n g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}, \quad c_n := \left(\frac{n-3}{n-1} - \frac{1}{4}\right). \tag{3.4}$$

The Maxwell equations are, J being the convection electric current of the density of charge q,

$$dF = 0 \quad \text{and} \quad \nabla . F = J, \quad \text{i.e.} \quad \nabla_\alpha F^{\alpha\beta} = J^\beta := qu^\beta; \tag{3.5}$$

they imply the conservation of charge

$$\nabla_\alpha (qu^\alpha) = 0. \tag{3.6}$$

One introduces the electromagnetic potential A, 1-form such that (for simplicity we consider a domain where A exists globally):

$$F = dA. \tag{3.7}$$

We take A in Lorentz gauge, i.e. such that

$$\delta A = 0. \tag{3.8}$$

The Maxwell equations read then as a wave equation for A, namely

$$\nabla_\alpha \partial^\alpha A^\beta - R^\beta{}_\lambda A^\lambda = J^\beta = q u^\beta, \tag{3.9}$$

where we can replace the Ricci tensor $R_{\alpha\beta}$ by its value (3.3).

Modulo the Maxwell equations (and $u^\alpha u_\alpha = -1$), the stress energy conservation equations are equivalent to

$$\nabla_\alpha (r u^\alpha) = 0 \tag{3.10}$$

and

$$r u^\alpha \nabla_\alpha u^\beta + q u_\lambda F^{\beta\lambda} = 0. \tag{3.11}$$

Equations (3.10) and (3.6) imply that

$$u^\alpha \partial_\alpha \frac{q}{r} = 0;$$

therefore q/r is constant along the flow lines, and it is constant throughout the spacetime we construct if constant initially. We will make this simplifying (though not necessary) hypothesis, and set

$$q = kr,$$

with k some given constant. Equation (3.11) can then be replaced by

$$u^\alpha \nabla_\alpha u^\beta + k u_\lambda F^{\beta\lambda} = 0. \tag{3.12}$$

3.2 *Existence and uniqueness theorem in wave and Lorentz gauges*

Theorem 3.1 *The Einstein equations in a wave gauge with sources an electromagnetic field of potential in a Lorentz gauge together with charged pure matter are a hyperbolic Leray system for g, A, q, u, and r, which is causal as long as g is Lorentzian and u is timelike.*

Corollary 3.2 *There is an interval $[0, T) \subset R$ such that the Cauchy problem for these equations with data on the manifold M given by $\bar{g} \in M_s$, $\bar{A} \in H_s, \bar{K}, \bar{F} \in H_{s-1}$, $\bar{v} \in H_{s-1}$, $|\bar{v}|_{\bar{g}} < 1$, $\bar{r} \in H_{s-2}$, $s > \frac{n}{2} + 2$, $\frac{\bar{q}}{\bar{r}} = k$, a constant, has one and only one solution $g \in L_s(T)$, $u \in U_{s-1}(T)$, $q, r \in E_{s-2}(T)$.*

Corollary 3.3 *If the initial data satisfy the Einstein and Maxwell constraints, the solution in wave and Lorentz gauges satisfies the original Einstein–Maxwell dust system. There is a unique maximal globally hyperbolic solution if $s > \frac{n}{2} + 3$.*

Proof As in the previous section we differentiate the Einstein equations in a wave gauge in the direction of u. We obtain, using (3.3)

$$u^\gamma \nabla_\gamma R^{(h)}_{\alpha\beta} = f_{\alpha\beta}, \tag{3.13}$$

with (we could replace $u^\gamma \nabla_\gamma u_\lambda$ by its value taken from (3.11), but it is not necessary for our results)

$$f_{\alpha\beta} := u^\gamma \nabla_\gamma \Phi_{\alpha\beta} - r\left\{\nabla_\gamma u^\gamma \left(u_\alpha u_\beta + \frac{1}{2} g_{\alpha\beta}\right) + u^\gamma \nabla_\gamma (u_\alpha u_\beta)\right\}. \tag{3.14}$$

We also differentiate Equations (3.9) for the potential in the direction of u. We obtain:

$$u^\gamma \nabla_\gamma \nabla_\alpha \partial^\alpha A^\beta - f^\beta{}_\lambda A^\lambda - R^\beta_\lambda u^\gamma \nabla_\gamma A_\lambda = kr\{-u^\beta \nabla_\gamma u^\gamma + u^\gamma \nabla_\gamma u^\beta\}, \tag{3.15}$$

The system is composed of Equations (3.13), (3.15), (3.12), and (3.10) for the unknowns g, A, u, r. We give to the equations and unknowns the Leray–Volevic indices

$$m(3.13) = m(3.15) = 0, \qquad m(3.12) = 1, \quad m(3.10) = 0, \tag{3.16}$$

and to the unknowns g, u, r the indices

$$\ell(g) = \ell(A) = 3, \quad \ell(u) = 2, \quad \ell(r) = 1. \tag{3.17}$$

We see that the matrix of principal parts is again diagonal, with the same kind of terms in the diagonal as in the previous section. The existence and uniqueness theorem in wave and Lorentz gauges follows from the Leray–Dionne theory.

The existence of a solution of the full Einstein–Maxwell dust system when the initial data satisfy the Einstein and Maxwell constraints follows the same lines as in the vacuum Einstein–Maxwell case. Geometric global uniqueness is also proved similarly. □

3.3 *Motion of isolated bodies*

Theorem 2.2 extends to the case of charged dust, because the flow lines of u depend only on g and F, while the right-hand sides of (3.14) and (3.15) vanish outside of the support of $\bar{r}$.

4 Perfect fluid, Euler equations

We wrote in Equation (II.8.2) the stress energy tensor of a perfect fluid in Special Relativity. Following the general principle given in the introduction of this chapter, the stress energy tensor T of a general relativistic perfect fluid is taken to be

$$T_{\alpha\beta} = \mu u_\alpha u_\beta + p(g_{\alpha\beta} + u_\alpha u_\beta), \tag{4.1}$$

where u is the unit flow vector, while μ and p are respectively the energy and pressure densities. Indeed, the components of T at a point in an orthonormal Lorentz frame with timelike vector u, called a **proper frame**, are μ for the time–time component and zero for the mixed time–space component, while the space components, representing the stresses due to pressure, are the isotropic space tensor $p\delta_{ij}$. For classical fluids μ and p are non-negative.

The conservation laws of a perfect fluid are

$$\nabla_\alpha T^{\alpha\beta} \equiv (\mu + p)u^\alpha \nabla_\alpha u^\beta + g^{\alpha\beta}\partial_\alpha p + u^\beta[\nabla_\alpha(\mu + p)u^\alpha] = 0.$$

By contracted product with u_β one deduces from these equations the energy equation

$$(\mu + p)\nabla_\alpha u^\alpha + u^\alpha \partial_\alpha \mu = 0, \tag{4.2}$$

and, using this equation (or contracting with the projection operator $\pi_{\alpha\beta} \equiv g_{\alpha\beta} + u_\alpha u_\beta$), the equations of motion:

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\alpha u^\beta)\partial_\alpha p = 0. \tag{4.3}$$

Equations (4.2) and (4.3) are called the **Euler equations**.

Remark 4.1 *The equations of motion imply $u^\alpha \partial_\alpha(u^\beta u_\beta) = 0$: the condition $u^\alpha u_\alpha = -1$ is conserved along the flow lines.*

Remark 4.2 *The trace of the stress energy tensor is, with n the dimension of space (classical case, $n = 3$):*

$$T \equiv g^{\alpha\beta}T_{\alpha\beta} = np - \mu; \tag{4.4}$$

The Einstein equations with source such a perfect fluid can be written as follows:

$$R_{\alpha\beta} = \rho_{\alpha\beta},$$

$$\rho_{\alpha\beta} := T_{\alpha\beta} - \frac{1}{n-1}g_{\alpha\beta}T = (\mu + p)u_\alpha u_\beta + \frac{1}{n-1}g_{\alpha\beta}(\mu - p). \tag{4.5}$$

5 Energy properties

We defined in Chapter 3 some energy conditions frequently referred to in the literature. For a perfect fluid one has the following propositions.

Proposition 5.1 *If μ and p are non-negative, condition satisfied in the absence of non classical effects, while X is causal then*

1. *The scalar $T^{\alpha\beta}X_\alpha X_\beta$ is non-negative; one says that a perfect fluid satisfies the* ***weak energy condition****.*
2. *The scalar $\rho_{\alpha\beta}X^\alpha X^\beta$ is non-negative such a perfect fluid satisfies the* ***strong energy condition****.*

Proof

1. It holds that

$$T^{\alpha\beta}X_\beta X_\alpha \equiv \mu(u^\alpha X_\alpha)(u^\beta X_\beta) + p\{(u^\alpha X_\alpha)(u^\beta X_\beta) + X^\alpha X_\alpha\}. \tag{5.1}$$

   We see immediately that this scalar is non-negative under the given conditions by computing its value in a proper frame.

2. In a proper frame the tensor $\rho_{\alpha\beta} \equiv T_{\alpha\beta} - \frac{1}{n-1} g_{\alpha\beta} T$ is such that:

$$\rho_{\alpha\beta} X^\alpha X^\beta = \frac{1}{n-1}\mu \left\{ (n-2)(X^0)^2 + \sum (X^i)^2 \right\} + \frac{\rho}{n-1} \left\{ \frac{3}{2}(X^0)^2 - \sum (X^i)^2 \right\}. \tag{5.2}$$

Therefore for a causal X, for which $(X^0)^2 \geq \sum (X^i)^2$:

$$\rho_{\alpha\beta} X^\alpha X^\beta \geq \left(\frac{n-2}{n-1}\mu + p \right) (X^0)^2. \tag{5.3}$$

□

The **energy momentum vector** relative to a timelike vector X is

$$P_X^\alpha := T^{\alpha\beta} X_\beta \equiv (\mu + p) u^\alpha u^\beta X_\beta + p X^\alpha. \tag{5.4}$$

Proposition 5.2 *If $\mu \geq |p|$ and if X is past timelike, the energy momentum vector P_X is future timelike; one says that such a perfect fluid satisfies the* ***dominant energy condition****.*

Proof In a proper frame the components of P_X are

$$P_X^0 = \mu X_0, \quad P_X^i = p X^i \tag{5.5}$$

□

6 Particle number conservation

In relativity, mass and energy are the same entity. However, Taub[6] pointed out the existence of another scalar function of physical interest, the particle number density r (better called the rest mass density if there are different kinds of particles with non-zero rest mass). In the absence of chemical reactions or quantum phenomena, conservation of particle number implies the equation

$$\nabla_\alpha (r u^\alpha) = 0. \tag{6.1}$$

The integration of this equation on a spacetime tube $W \subset V = M \times R$ with boundaries $W_{t_0} \subset M_{t_0}$, $W_{t_1} \subset M_{t_1}$ and a lateral boundary L generated by flow lines, gives, using Stokes' formula,

$$\int_W \nabla_\alpha (r u^\alpha) \mu_g = 0 = \int_{\partial W} r u^\alpha \nu_\alpha \mu_{\partial W} = \int_{W_{t_1}} r u^0 N \mu_{\bar{g}} - \int_{W_{t_0}} r u^0 N \mu_{\bar{g}}, \tag{6.2}$$

where ν is the unit normal to the boundary ∂W, N denotes the lapse of the foliation $M \times R$ and $g_{\partial W}$ the induced metric on ∂W, and $\bar{g}$ is the induced metric on a leaf M_t. We deduce from this formula that the t-dependent space scalar

$$\bar{r} \equiv r u^0 N \tag{6.3}$$

represents the density of flow lines crossing a submanifold W_t.

The quantity r^{-1} plays the role of a specific volume.

[6] Taub, A. H. (1959) *Arch. Rat. Mech. An.*, **3**, 312–29.

7 Thermodynamics

7.1 *Definitions. Conservation of entropy*

The difference between the total energy density μ and the rest mass r density is called the **internal energy** density. One denotes by ε the specific internal energy density; that is, one sets

$$\mu = r(1 + \varepsilon). \tag{7.1}$$

In the case of local thermodynamic equilibrium (reversible thermodynamics) one defines a specific entropy density S and an absolute temperature T by extending to relativistic perfect fluids the identity of the first law of thermodynamics, namely:

$$TdS := d\varepsilon + pd(r^{-1}). \tag{7.2}$$

The thermodynamic quantities μ, p, S, T are spacetime scalar functions.

Theorem 7.1 *(Pichon)* [7] *In a perfect fluid the thermodynamic identity (7.2) and the matter conservation equation (7.1) imply the conservation of the entropy density along the flow lines:*

$$u^\alpha \partial_\alpha S = 0 \tag{7.3}$$

Proof The identity (7.2) and the definition (7.1) of ε give that

$$Tu^\alpha \partial_\alpha S \equiv u^\alpha \partial_\alpha \varepsilon - r^{-2} p u^\alpha \partial_\alpha r \equiv u^\alpha r^{-1} \partial_\alpha \mu - (\mu + p) r^{-2} u^\alpha \partial_\alpha r;$$

hence, using the energy equation (4.2),

$$Tu^\alpha \partial_\alpha S \equiv -r^{-2}(\mu + p) \nabla_\alpha (r u^\alpha). \qquad \square$$

7.2 *Equations of state*

For a perfect fluid only two thermodynamic scalars are independent; the others are linked to them by relations which are assumed to depend only on the nature of the given fluid. A usual general formula for the equation of state is to give p as some smooth function of μ and S,

$$p = p(S, \mu),$$

invertible as a smooth function $\mu = \mu(p, S)$.

Two circumstances are of particular interest in General Relativity: astrophysics and cosmology.

[7] Pichon, G. (1965) *Ann. Inst. Henri Poincaré A*, **2**, 21–85.

7.2.1 *Astrophysics*

Here one is inspired by what is known from classical fluids, with additional relativistic considerations. Particularly interesting cases are:

- Barotropic fluids. There the equation of state reduces to

$$p = p(\mu)$$

The fluid dynamics are then governed by the energy and momentum equations. The particle number conservation equation decouples from the others and can be solved after the fluid motion has been determined. Some physical situations which correspond to this modelling are the following:

1. Very cold matter, models of nuclear matter.
2. Ultrarelativistic fluids, i.e. fluids in thermal equilibrium, where the energy μ is largely dominated by the radiation energy. Then, by the Stefan–Boltzmann laws, it holds that $\mu = KT^4$, $p = \frac{1}{3}KT^4$, and hence

$$p = \frac{1}{3}\mu \tag{7.4}$$

The same equations for p and μ (with different constants K) and the equation of state (7.4) are valid for a fluid of massless neutrinos or electron–positron pairs.

Remark 7.2 *The stress energy tensor of an ultrarelativistic fluid is traceless.*

- Polytropic fluids. They obey, like their classical analogues, an equation of state of the form

$$p = f(S)r^\gamma$$

Several physical situations correspond to this case.

Average stellar situations: there only the internal energy ε and pressure p are dominated by radiation, then $\varepsilon = KrT^4$ and

$$p = \frac{1}{3}KT^4 \quad \text{hence} \quad p = \frac{1}{3}r\varepsilon \tag{7.5}$$

On the other hand the thermodynamic identity together with the expressions of ε and p implies

$$dS = \frac{4K}{3}d(r^{-1}T^3) \quad \text{hence} \quad S = \frac{4KT^3}{3r} \tag{7.6}$$

eliminating T between (7.5) and (7.6) gives the polytropic equation of state of index $\gamma = \frac{4}{3}$:

$$p = \frac{K}{3}\left(\frac{3S}{4K}\right)^{\frac{4}{3}} r^{\frac{4}{3}}, \quad \text{with} \quad \mu = 3p + r. \tag{7.7}$$

- More refined equations of state adapted to various physical situations have been considered.

7.2.2 *Cosmology*

There is little physical information about the fluid which is present in cosmological models.

It is assumed that in the early Universe of the Big Bang models, at very high temperature the fluid was ultrarelativistic, i.e. obeyed the equation of state (7.4).

At later times, after the formation of galaxies, it is in general assumed, for simplicity, that there is a linear equation of state independent of entropy

$$p = (\gamma - 1)\mu. \tag{7.8}$$

In order that the speed of sound waves not be greater than the speed of light one supposes that $1 \leq \gamma \leq 2$. The case $\gamma = 1$ corresponds to dust. In the case $\gamma = 2$ the fluid is called stiff, or incompressible. In a stiff fluid the sound and light speeds are equal, as we will see in the following section.

8 Wave fronts, propagation speeds, shocks

In Newtonian mechanics a wave front is a 2-surface in space which propagates with time. Its propagation speed at a point of space and an instant of time is the quotient by an infinitesimal absolute time interval, δt, of the infinitesimal distance between two wave fronts measured in the direction orthogonal to them at times t and $t+\delta t$. In relativity, a wave front is a 3-submanifold of the spacetime. The definition of its propagation speed depends on the choice of observer, and requires some thought.

8.1 *General definitions*

The wave fronts associated with a system of partial differential equations on a spacetime are submanifolds of this spacetime whose normals are roots of the characteristic determinant[8]. In other words, wave fronts are given by equations $f = 0$, where the scalar function f a is solution of the first-order partial differential equation, called the **eikonal** equation:

$$\Phi\left(x, \frac{\partial f}{\partial x}\right) = 0. \tag{8.1}$$

The eikonal equation is constructed by replacing in the characteristic determinant the covector X by the partial derivative ∂, and equating to zero the action on f of the obtained operator. Generically, discontinuities of the derivatives of the unknowns of the order appearing in these principal parts can occur only across such submanifolds.

[8] See the section on Leray hyperbolic systems in Appendix IV.

A **wave front** is generated[9] by the bicharacteristics of this equation, also called rays, which are solutions of the ordinary differential system (one sets $p = \partial f/\partial x$ in (8.1), λ is a parameter, called canonical)

$$\frac{dx^\alpha}{\partial\Phi/p_\alpha} = -\frac{dp_\alpha}{\partial_\alpha\Phi/\partial x^\alpha} = d\lambda. \tag{8.2}$$

This system admits the first integral

$$\Phi(x,p) = constant. \tag{8.3}$$

The tangents to the rays issued from some given point x generate a cone in the tangent space to the spacetime at x, called the wave cone, dual to the characteristic cone. The wave cone is the envelope of the hyperplanes whose normals (in the metric g) belong to the characteristic cone. A wave front at x is tangent to the wave cone at x along the direction of a ray.

As an example we prove the following lemma, stating the fact that null rays are orthogonal (in the metric g) to null surfaces.

Lemma 8.1 *The light cone of a spacetime (V, g) and its dual can be identified by the usual identification of the tangent and cotangent space to V through the metric g.*

Proof The eikonal equation associated with the light cone of g is

$$g^{\alpha\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta} = 0. \tag{8.4}$$

The light rays issued from some given point satisfy the equations

$$v^\alpha := \frac{dx^\alpha}{d\lambda} = 2g^{\alpha\beta}p_\beta, \qquad g^{\alpha\beta}p_\alpha p_\beta = 0, \tag{8.5}$$

hence, using the identification through g of covariant and contravariant tensors, the equation of the light wave cone: $g_{\alpha\beta}v^\alpha v^\beta = 0$. □

The propagation speed of a smooth wave front with respect to an observer is the propagation speed of the tangent plane to the wave front with respect to the proper Lorentz frame of this observer, with the following definition.

Definition 8.2 *The **propagation speed** with respect to an orthonormal Lorentzian frame at a point x of a three-dimensional*[10] *hyperplane P is the velocity with respect to this frame of the vector in P of greatest slope with respect to the space hyperplane of the frame. The greatest slope of an hyperplane P with respect to an orthonormal frame is a vector in P orthogonal to the 2-plane*[11] *intersection of the space hyperplane, $X^0 = 0$, and P.*

[9] See for instance CB-DM1, IV C 7.

[10] One could give analogous definitions for higher dimensions, but to be clearer, we prefer to stick to $3 + 1$, which is the only one of practical use for classical fluids.

[11] Two 3-hyperplanes in four-dimensional spacetimes intersect generically along a 2-plane.

Lemma 8.3 *The propagation speed V of a hyperplane with normal ν with respect to a Lorentzian frame with time vector u is given by the following formula*

$$|V|^2 = \frac{(u^\alpha \nu_\alpha)^2}{(u^\alpha u^\beta + g^{\alpha\beta})\nu_\alpha \nu_\beta}. \tag{8.6}$$

Proof The vector of greatest slope in P with respect to the Lorentz frame with time vector u is the vector in P orthogonal to the 2-plane I which is the intersection of P and the hyperplane orthogonal to u. This intersection satisfies the equations

$$u^\alpha X_\alpha = 0, \quad \nu^\alpha X_\alpha = 0. \tag{8.7}$$

Choose for the first space vector in the Lorentzian frame the projection of ν on the hyperplane orthogonal to u, the above equations read:

$$X_0 = 0, \quad \nu^0 X_0 + \nu^1 X_1 = 0. \tag{8.8}$$

A vector in I has components $X_0 = 0$, $X_1 = 0$, X_2, X_3 arbitrary. The vector of greatest slope Y in P is uniquely determined by the conditions

$$Y \in P, \quad \text{i.e.} \quad \nu^0 Y_0 + \nu^1 Y_1 = 0, \quad \text{hence} \quad Y_1 = -\frac{\nu^0}{\nu^1} Y_0, \tag{8.9}$$

and the orthogonality with I, that is $Y_2 = Y_3 = 0$.

The propagation speed of P with respect to the considered Lorentz frame is

$$|V| = \left|\frac{Y_1}{Y_0}\right| = \left|\frac{\nu^0}{\nu^1}\right| \tag{8.10}$$

The given formula (8.10) takes the form (8.6) in a Lorentz frame where the time axis is the unit vector u and the normal ν has components $\nu_2 = \nu_3 = 0$. □

8.2 *Case of perfect fluids*

The Euler and entropy conservation equations of a perfect fluid are of first order. Using the entropy equation (7.3) in the energy equation (4.2) we see that the Euler-entropy equations written for the unknowns S, p, u have a characteristic matrix composed of two blocks around the diagonal. One corresponds to S and the conservation of entropy (7.3); it reduces to

$$a \equiv u^\alpha X_\alpha.$$

The other block corresponds to the unknowns p and u and Equations (4.2) and (4.3). It is the following 5×5 matrix[12] with $\rho \equiv \mu + p$, $\mu'_p = \partial\mu/\partial p$

$$\mathcal{M} \equiv \begin{pmatrix} a\mu'_p & \rho X_0 & \rho X_1 & \rho X_2 & \rho X_3 \\ X^0 + au^0 & \rho a & 0 & 0 & 0 \\ X^1 + au^1 & 0 & \rho a & 0 & 0 \\ X^2 + au^2 & 0 & 0 & \rho a & 0 \\ X^3 + au^3 & 0 & 0 & 0 & \rho a \end{pmatrix}$$

The determinant of this matrix is computed to be:

$$-\rho^4 a^3 D \tag{8.11}$$

with

$$D \equiv (-\mu'_p + 1)(u^\alpha X_\alpha)^2 + X^\alpha X_\alpha \tag{8.12}$$

We see that a perfect fluid has two types of wave front:

- The matter wave fronts, $f = constant$, such that

$$u^\alpha \partial_\alpha f = 0. \tag{8.13}$$

  These are submanifolds generated by the flow lines. Their propagation speed for a comoving observer (i.e. in a proper rest frame of the fluid) is zero.
- The sound wave fronts such that their normals satisfy $D = 0$. In a proper rest frame $g_{\alpha\beta} = \eta_{\alpha\beta}$ (the Minkowski metric), $u^0 = 1, u^i = 0$. The corresponding eikonal equation reads:

$$-\mu'_p(\partial_0 f)^2 + \sum_i (\partial_i f)^2 = 0, \tag{8.14}$$

  the propagation speed of these wave fronts is, assuming $\mu'_p > 0$,

$$|V| = (\mu'_p)^{-\frac{1}{2}};$$

  it is less than the speed of light, as expected from a relativistic theory, if and only if:

$$\mu'_p \geq 1. \tag{8.15}$$

  A fluid such that $\mu'_p = 1$ is called an **incompressible** or **stiff** fluid: in such a fluid the sound waves propagate with the speed of light.

8.3 *Shocks*

A shock is a discontinuity in the fluid variables across a timelike n-manifold $n = 3$ in the classical case. The stress energy tensor is then discontinuous; its derivative is meaningful only in a generalized sense.

[12] The four components of u are considered as independent unknowns. The identity $g(u, u) = -1$ is preserved by the flow.

The relativistic **Rankine–Hugoniot equations** express the vanishing of the divergence of the stress energy tensor in the space of generalized functions (distributions); they are

$$n_\alpha[T^{\alpha\beta}] = 0,$$

where n is the spacelike normal to the timelike shock front Σ and $[T^{\alpha\beta}]$ is a measure with spport Σ. A deep study of the formation of shocks in relativistic fluids has been done by Christodoulou[13].

9 Stationary motion

A fluid is said to be in stationary motion if u, r, and the thermodynamic scalars are invariant under a 1-parameter isometry group whose trajectories are timelike. Taking these trajectories, which do not coincide in general with the flow lines, as time lines, the time derivatives in the Euler equations drop out and the determinant D reduces to (note that $p'_\mu \equiv 1/\mu'_p$)

$$D \equiv (p'_\mu - 1)(u^i X_i)^2 + g^{ij} X_i X_j \tag{9.1}$$

Keeping the Killing vector as collinear to the time vector e_0 of a Lorentzian orthonormal frame (called then a stationary frame) but taking as its first space axis the projection of u on the hyperplane orthogonal to e_0, then $u^i = 0$ if $i \neq 1$ and D reads

$$D \equiv (X_1)^2\{(u^1)^2(p'_\mu - 1)\} + p'_\mu \sum_{i\geq 1} (X_i)^2. \tag{9.2}$$

This quadratic form is positive if

$$(u^1)^2 < \frac{p'_\mu}{1 - p'_\mu} \tag{9.3}$$

and hyperbolic for the reverse inequality. The velocity of the flow in the stationary frame is

$$v = \frac{u^1}{u^0} = \frac{u^1}{\sqrt{1 + (u^1)^2}} = \sqrt{p'_\mu} \tag{9.4}$$

The quadratic form D is therefore positive when $v < \sqrt{p'_\mu}$, hyperbolic for the reverse inequality. We see that, as in classical mechanics, stationary flows with subsonic speed satisfy elliptic equations, while supersonic flows satisfy hyperbolic equations.

10 Dynamic velocity for barotropic fluids

10.1 *Fluid index and equations*

Important properties of non-relativistic fluids generalize to relativistic fluids if one introduces a spacetime vector linked with both the kinematic unit vector u and the thermodynamic quantities.

[13] Christodoulou, D. (2007) *The Formation of Shocks in Relativistic Fluids*, EMS, Zurich.

In the case of barotropic fluids, i.e. with equation of state $\mu = \mu(p)$, the simplest way is to define a function, $f(p)$, called the index of the fluid, by

$$f(p) := \exp \int \frac{dp}{\mu(p) + p} \tag{10.1}$$

and the dynamic 4-velocity by

$$C_\alpha = f u_\alpha, \quad \text{hence} \quad C^\alpha C_\alpha = -f^2$$

It holds that

$$\partial_\alpha f \equiv \frac{\partial f}{\partial x^\alpha} = f^{-1} \frac{\partial_\alpha p}{\mu + p} \tag{10.2}$$

also

$$\partial_\alpha f = -f^{-1} C^\beta \nabla_\alpha C_\beta \tag{10.3}$$

Theorem 10.1

1. *For a barotropic fluid the Euler equations (4.2) and (4.3) are equivalent to the following equations:*

$$C^\alpha (\nabla_\alpha C_\beta - \nabla_\beta C_\alpha) = 0 \tag{10.4}$$

and

$$\nabla_\alpha C^\alpha + (\mu'_p - 1) \frac{C^\alpha C^\beta}{C^\lambda C_\lambda} \nabla_\alpha C_\beta = 0. \tag{10.5}$$

We have denoted $\mu'_p \equiv \frac{\partial \mu}{\partial p}$. which is a known function of p when the equation of state is known. If one expresses p as a function of f the perfect fluid equations read as a first-order differential system for C.

2. *The equations (10.4) express that the flow lines are geodesics of the following metric $\tilde{g}$ conformal to the spacetime metric g:*

$$\tilde{g} := f^2 g. \tag{10.6}$$

Proof

1. We have

$$u^\alpha \nabla_\alpha u_\beta \equiv f^{-1} C^\alpha \nabla_\alpha (f^{-1} C_\beta) \equiv f^{-2} (C^\alpha \nabla_\alpha C_\beta - f^{-1} C^\alpha C_\beta \partial_\alpha f) \tag{10.7}$$

and

$$(g^\alpha_\beta + u^\alpha u_\beta) \frac{\partial_\alpha p}{\mu + p} \equiv f \partial_\beta f + f^{-1} C^\alpha C_\beta \partial_\alpha f. \tag{10.8}$$

Therefore the identity for the equations of motion

$$u^\alpha \nabla_\alpha u_\beta + (g^\alpha_\beta + u^\alpha u_\beta) \frac{\partial_\alpha p}{\mu + p} \equiv f^{-2} [C^\alpha \nabla_\alpha C_\beta - C^\alpha \nabla_\beta C_\alpha]. \tag{10.9}$$

The other equation is deduced from the energy equation by an analogous computation.

2. If $\tilde{g} = f^2 g$ it holds that

$$\tilde{\nabla}_\beta C_\alpha - \nabla_\beta C_\alpha \equiv (\Gamma^\lambda_{\beta\alpha} - \tilde{\Gamma}^\lambda_{\beta\alpha}) C_\lambda \equiv -f^{-3}\{\partial_\alpha f C_\beta + \partial_\beta f C_\alpha - g_{\alpha\beta} g^{\lambda\mu} \partial_\mu f C_\lambda\}, \tag{10.10}$$

hence, using (10.3) and (10.4)

$$C^\beta \tilde{\nabla}_\beta C_\alpha \equiv C^\beta \nabla_\alpha C_\beta - f^{-2} \partial_\alpha f C^\beta C_\beta = 0. \tag{10.11}$$

□

Remark 10.2 *We see from Equation (10.5) that a barotropic relativistic fluid is incompressible ($\mu'_p = 1$) if and only if*

$$\nabla_\alpha C^\alpha = 0. \tag{10.12}$$

This incompressibility property generalizes the classical one for Newtonian fluids, $\partial_i v^i = 0$, implied in this case by the constancy of density.

10.2 *Vorticity tensor and Helmholtz equations*

The vorticity tensor is defined through the dynamic velocity as the antisymmetric 2-tensor

$$\Omega_{\alpha\beta} \equiv \nabla_\alpha C_\beta - \nabla_\beta C_\alpha.$$

It results from the equations of motion (10.4) that this tensor is orthogonal to the velocity, since these equations read

$$C^\alpha \Omega_{\alpha\beta} = 0.$$

Theorem 10.3 *(Relativistic Helmholtz equations) The Lie derivative of the vorticity tensor Ω with respect to the dynamic velocity C vanishes:*

$$\mathcal{L}_C \Omega = 0.$$

Proof The derivation of (10.4) in the direction of C gives

$$C^\alpha \nabla_\gamma \Omega_{\alpha\beta} + \nabla_\gamma C^\alpha \Omega_{\alpha\beta} = 0.$$

On the other hand, since Ω is an exterior 2-form differential of the 1-form C, it is a closed form, satisfying in local coordinates the identity

$$\nabla_\alpha \Omega_{\beta\gamma} + \nabla_\gamma \Omega_{\alpha\beta} + \nabla_\beta \Omega_{\gamma\alpha} \equiv 0.$$

Combining these two relations gives the announced equation:

$$C^\alpha \nabla_\alpha \Omega_{\beta\gamma} + \nabla_\gamma C^\alpha \Omega_{\beta\alpha} + \nabla_\beta C^\alpha \Omega_{\alpha\gamma} = 0. \tag{10.13}$$

□

We have shown that the vorticity tensor satisfies a linear differential homogeneous system along the flow lines, hence the following corollary.

Corollary 10.4 *If a smooth (barotropic) flow has vanishing vorticity on a 3-submanifold transversal to the flow lines it has a vanishing vorticity on the domain of spacetime spanned by these flow lines.*

10.3 *Irrotational flows*

10.3.1 *Fluid equations*

A flow with zero vorticity is called irrotational. Its trajectories are locally orthogonal to hypersurfaces, since the equation $\nabla_\alpha C_\beta - \nabla_\beta C_\alpha = 0$ implies that there exists on spacetime, at least locally, a function Φ such that

$$C_\alpha = \partial_\alpha \Phi.$$

One can prove the preservation of irrotationality using the fact that the flow lines are geodesics of the conformal metric (10.6).

Lemma 10.5 *The irrotational flow of a perfect fluid is governed by a quasilinear wave type equation*[14].

Proof When the flow is irrotational Equation (10.4) is identically satisfied while (10.5) reads:

$$(g^{\alpha\beta} + (1 - \mu'_p)u^\alpha u^\beta)\nabla_\alpha \partial_\beta \Phi = 0 \tag{10.14}$$

In this equation (10.14) the quantities u and p are given functions of Φ and $\partial\Phi$. Indeed, by definition

$$g^{\alpha\beta}\partial_\alpha\Phi\partial_\beta\Phi \equiv -f^2 \quad \text{and} \quad u^\alpha \equiv \frac{\partial^\alpha \Phi}{f}$$

while p can be expressed in terms of f (i.e. of $\partial\Phi$) by inverting the relation $C^\alpha C_\alpha = -f^2$.

The characteristic cone of this quasilinear second-order differential equation for Φ, is given by

$$(g^{\alpha\beta} + (1 - \mu'_p)u^\alpha u^\beta)X_\alpha X_\beta = 0$$

Since u is a timelike unit vector for the metric g the quadratic form above is of Lorentzian signature if and only if $\mu'_p > 0$. Equation (10.14) is then hyperbolic. The dual of its characteristic cone is called the sound cone. It is causal (i.e. interior to the light cone) if $\mu'_p \geq 1$. □

Remark 10.6 *In the case of a stiff fluid, $\mu'_p = 1$ the equation for Φ reduces to the usual linear wave equation.*

Example 10.7 *Take the cosmological equation of state $p = (\gamma - 1)\mu$. Then f is given by*

$$f \equiv \exp \int \frac{(\gamma - 1)dp}{\gamma p}$$

hence, up to an irrelevant multiplicative constant,

$$f \equiv p^{\frac{\gamma-1}{\gamma}}, \quad p = f^{\frac{\gamma}{\gamma-1}}$$

[14] Fourès (Choquet)-Bruhat, Y. (1958) *Bull. Soc. Math. France*, **86**, 155–75.

10.3.2 *Coupling with Einstein equations*

The Einstein equations in a harmonic gauge with source given by an irrotational flow, together with the equation for Φ, form a second-order quasidiagonal system whose principal parts are the wave operator of the spacetime metric or the sound wave operator. It is a Leray hyperbolic system as soon as $\mu'_p > 0$. However, the system is causal (i.e. the domain of dependence is determined by the light cone) only if $\mu'_p \geq 1$.

Remark 10.8 *The solution can be interpreted as an Einsteinian spacetime with source an irrotational flow as long as $\partial\Phi$ is timelike, i.e. $g^{\alpha\beta}\partial_\alpha\Phi\partial_\beta\Phi < 0$.*

Remark 10.9 *A submanifold which is spacelike for the spacetime metric gives rise to a well-posed Cauchy problem only if its tangent plane is also exterior to the sound cone. This is always the case if $\mu'_p \geq 1$.*

Exercise. *Formulate the Cauchy problem and a local existence theorem for the Einstein equations coupled with a perfect fluid irrotational flow.*

11 General perfect fluids

For non-barotropic fluids, one introduces the enthalpy h, defined through the thermodynamic identity

$$dh \equiv r^{-1}dp + TdS, \tag{11.1}$$

or equivalently, due to the first thermodynamic identity (7.2)

$$dh \equiv d[r^{-1}(\mu + p)], \quad \text{we set} \quad h := r^{-1}(\mu + p), \tag{11.2}$$

and define now the dynamic velocity by

$$C^\alpha := hu^\alpha, \tag{11.3}$$

hence

$$C^\alpha \nabla_\beta C_\alpha = -h\partial_\beta h. \tag{11.4}$$

A straightforward calculation shows that the equations of motion (4.3) are equivalent to

$$(hT)^{-1}C^\alpha \Omega_{\alpha\beta} = \partial_\beta S, \tag{11.5}$$

where, with the definition (11.3) of C^α,

$$\Omega_{\alpha\beta} := (\nabla_\alpha C_\beta - \nabla_\beta C_\alpha). \tag{11.6}$$

The fact that Ω with its new definition is again a closed form, since the right-hand side of (11.6) is an exact differential, permits the extension of the corollary of Theorem 10.3 to these more general fluids, by showing that Ω still satisfies a linear homogeneous first-order differential equation,

$$(hT)^{-1}C^\alpha \nabla_\alpha \Omega_{\beta\gamma} + \nabla_\gamma[(hT)^{-1}C^\alpha]\Omega_{\beta\alpha} + \nabla_\beta[(hT)^{-1}C^\alpha]\Omega_{\alpha\gamma} = 0. \tag{11.7}$$

The energy equation reads, modulo the conservation of entropy along the flow lines,

$$(\mu + p)\nabla_\alpha C^\alpha + C^\alpha(\mu'_p - 1)\partial_\alpha p = 0,$$

We have again $\nabla_\alpha C^\alpha = 0$ when $\mu'_p = 1$ (incompressible fluids).

12 Hyperbolic Leray system

12.1 *Hyperbolicity of the Euler equations.*

The first proof of hyperbolicity of the Einstein Euler system was obtained[15] by coupling the Helmholtz equations with a quasidiagonal second-order system linking the dynamical velocity with the vorticity tensor. The principal operator in this system is the wave operator corresponding to the propagation of sound waves with speed $(\mu'_p)^{-\frac{1}{2}}$. It was proved that the complete system obtained by adding to the equations for Ω and C the Einstein equations in a wave gauge is a hyperbolic Leray system, which hence admits a well-posed Cauchy problem. The conservation in time of the Einstein constraints was checked to prove the existence of a solution of the full Einstein Euler system for initial data satisfying these constraints. The proof was extended to include the entropy law by Lichnerowicz (1967)[16]. We will not detail this proof here but give a direct proof using general results on quasidiagonalization of differential systems[17]. For simplicity we consider the classical case $n = 3$.

The considered equations are the Euler equations (4.2), (4.3) and the entropy conservation (7.3) for the six unknowns u^α, μ, S. We have seen that they imply

$$u^\alpha \partial_\alpha(u^\beta u_\beta + 1) = 0. \tag{12.1}$$

We have seen in Section 8.2 that the characteristic matrix is diagonal by blocks, one reducing to $a := u^\alpha X_\alpha$, the other denoted $\mathcal{M}$ corresponding to the Euler equations.

The determinant of $\mathcal{M}$ is, up to a numerical factor, $a^3 D$. The equation

$$a \equiv u^\alpha X_\alpha = 0 \tag{12.2}$$

determines a 3-plane orthogonal to u, while the cone:

$$D \equiv \{(\mu'_p - 1)u^\lambda u^\beta - g^{\lambda\beta}\} X_\lambda X_\beta = 0 \tag{12.3}$$

is a convex second-order cone. The product aD is an hyperbolic polynomial as soon as u is timelike and $p'_\mu > 0$. Its domain of dependence is determined by the sound cone, dual of the cone $D = 0$. Indeed if u is timelike the 3-plane orthogonal to it is exterior to the cone of normals to the sound cone: their equations in a proper frame are respectively $X^0 = 0$ and $-(X^0)^2 + p'_\mu \sum (X^i)^2 = 0$, therefore

[15] Choquet-Bruhat, Y. (1958) *Bull. Soc. Math.*, **86**, 155–75.

[16] Lichnerowicz, A. (1967) *Relativistic Hydrodynamics and Magnetohydrodynamics*, Benjamin, New York.

[17] Choquet-Bruhat, Y. (1966) *J. Math. Pures Appl.*, **45**, 371–86.

every straight line (in the considered tangent space) passing through a point inside the cone of normals to the sound cone cuts the third-order cone $aD = 0$ in three real distinct points. The Euler equations appear at first sight as a system which is hyperbolic only in the Leray–Ohya sense only, because the characteristic determinant of $\mathcal{M}$, a^3D, contains a multiple factor, a. However, it is easy to check that all cofactors of the matrix $\mathcal{M}$ can be divided by a^2. The general procedure of quasidiagonalization[18] shows that the Euler equations are equivalent to a quasidiagonal system with principal terms in the diagonal either $u^\alpha \partial_\alpha$ or

$$u^\alpha \{(\mu'_p - 1)u^\lambda u^\beta - g^{\lambda\beta}\}\partial_\lambda \partial_\beta \partial_\alpha. \tag{12.4}$$

Such a system satisfies the criteria of (strict) Leray hyperbolicity. We call the third-order system that we have constructed the *reduced Euler system.*

12.2 *Reduced Einstein–Euler entropy system*

One chooses a gauge such that the vacuum Einstein equations are well posed, for instance some (generalized or not) wave gauge. We consider the system of the reduced Einstein equations with fluid sources

$$R^{(h)}_{\alpha\beta} = (\mu + p)u_\alpha u_\beta + \frac{1}{2}(\mu - p)g_{\alpha\beta} \tag{12.5}$$

The quasidiagonal third-order system that we have obtained for the fluid variables contains derivatives of g up to third order since it is obtained by two derivations form the original system which contained first derivatives of g.

We choose the Leray indices as follows:

$$n(g) = 2, m(einstein) = 0, \quad n(u, p) = 1, m(fluid) = -2$$

The principal parts, elements of the characteristic matrix of the coupled system, are then:

- In the reduced Einstein equations (12.5) the principal parts for g are of order 2, and constitute a 10×10 diagonal matrix with elements (up to a numerical factor)
$$g^{\alpha\beta}\partial_\alpha \partial_\beta$$
The principal parts for u, μ, S are of order 1, and hence identically zero.
- In the reduced fluid equations the principal parts for u, μ, S are of order 3, they constitute a diagonal 5×5 matrix with elements given by (12.4).

The principal parts for g are of order 4, hence identically zero.

The light cone is interior to the characteristic cone of the coupled system, hence exterior to its wave cone if $\mu'_p \geq 1$.

We have proved the following theorem.

Theorem 12.1 *The reduced Einstein–Euler entropy system is a quasidiagonal Leray hyperbolic and causal system if* $\mu'_p \geq 1$.

[18] See the sections on Leray and Leray–Ohya hyperbolic systems in Appendix IV.

12.3 *Cauchy problem for the Einstein–Euler entropy system*

An initial data set for the Einstein–Euler system with a given equation of state will be the usual data for the Einstein equations, and in addition data for the fluid source.

We have seen that an initial data set for the Einstein equations is a triple $(M, \bar{g}, K)$ with M a 3-manifold, $\bar{g}$ a properly Riemannian metric and K a symmetric 2-tensor on M. A spacetime (V, g) is said to take these initial data if M can be embedded in V as a submanifold M_0 with induced metric $\bar{g}$ and extrinsic curvature K. If (V, g) is a solution of the Einstein equations with source the stress energy tensor T, then $(\bar{g}, K)$ must satisfy the following constraints

$$R(\bar{g}) - K.K + (trK)^2 = 2\rho, \tag{12.6}$$

$$\bar{\nabla}.K - \bar{\nabla} trK = J, \tag{12.7}$$

where ρ is a scalar and J a vector on $M\tilde{-}M_0$. In a Cauchy adapted frame, where the equation of M_0 in V is $x^0 \equiv t = 0$ and the time axis is orthogonal to M_0, one has:

$$\rho = N^2 T^{00}, \quad J^i = NT^{i0}$$

that is, in the case of a perfect fluid:

$$\rho = N^2(\mu + p)(u^0)^2 - p, \quad J^i = N(\mu + p)u^i u^0, \tag{12.8}$$

with, using $u^\alpha u_\alpha = -1$,

$$Nu^0 = (1 + g_{ij}u^i u^j)^{\frac{1}{2}}. \tag{12.9}$$

We see that on M_0 the quantities ρ and J depend only on the values $\bar{\mu}$ and $\bar{p}$ of μ and p and on the projection v of u, with components $v^i = u^i$, on M_0, they do not depend on the choice of lapse and shift. If the equation of state of the fluid is of the form $\mu = \mu(p, S)$ the original initial data for the fluid will be two scalars $\bar{p}$ and $\bar{S}$ on M and a tangent vector $\bar{v}$. Data for the third-order system are obtained by using the restriction to M_0 of the Euler equations and its first-order derivative in the direction of u. This computation requires now a choice for the initial lapse and shift, which is also required for the solution of the reduced Einstein equations.

We are ready to prove the following theorem.

Theorem 12.2 *Let $(M, \bar{g}, K, \bar{p}, \bar{S}, \bar{v})$ be an initial data set for the Einstein–Euler entropy equations with smooth equation of state $\mu(p, S)$. Suppose that $\bar{g} \in M_s$, $\partial\bar{g}, K, \bar{p}, \bar{S}, v \in H_{s-1}$, $s > \frac{3}{2} + 2$, and $\mu'_p(\bar{\mu}, \bar{S}) \geq 1$. Then there exists an Einsteinian spacetime (V, g) with source a perfect fluid with elements (u, p, S), depending continuously on the data, where (V, g) is a development of the initial data set $(M, \bar{g}, K)$ while v is the projection on M_0 of the flow vector u, and the energy density μ and entropy S reduce respectively on M_0 to $\bar{p}$ and $\bar{S}$.*

Proof We deduce from the initial value $v^i = \bar{u}^i$ the initial value for $\bar{u}^0$, after choosing an initial lapse $\bar{N}$, by using the condition $\bar{u}^\alpha \bar{u}_\alpha = -1$. We know that the reduced Einstein equations coupled with the reduced Euler entropy system form a Leray hyperbolic system for (g, u, p, S). For such a system the Cauchy problem is well posed in Sobolev spaces. The quasidiagonalization theory says that the solution, with Cauchy data determined as explained before, satisfies the original Euler entropy system. Equation (11.1) shows that the solution satisfies $u^\alpha u_\alpha = -1$. The stress energy tensor therefore satisfies the conservation laws.

The fact that the solution of the reduced system satisfies the full Einstein–Euler entropy equations follows from the general theorem in Chapter 6. □

Corollary 12.3 *The domain of dependence is determined by the light cone.*

Proof The light cone is exterior to the flow vector u and to the sound cone if $\mu'_p \geq 1$. □

12.4 *Motion of isolated bodies*

It has been pointed out by Rendall (1992)[19] that for a generic equation of state the theorem obtained above does not apply when there are domains of the initial manifold M where $\mu'_p(\bar{p}, \bar{S}) = 0$.

13 First-order symmetric hyperbolic system

Writing the Euler equations as a first-order symmetric hyperbolic system does not improve the results obtained by the Leray theory and destroys their space-time character, but may be preferred for some numerical computation, and also, in some cases, for the study of motion of isolated bodies, including regions where $\mu = 0$. To obtain such a first-order symmetric hyperbolic system one can apply general methods inaugurated by Lax, developed by Boillat and Ruggeri, and explained in the book by Anile (1987)[20]. These authors use a convex functional and auxiliary unknowns. We give here a direct computation in a sliced space-time, taking as unknowns p and the space components u^i, first without choosing a special frame. The component u^0 is determined through the identity

$$u^\alpha u_\alpha = -1. \tag{13.1}$$

The first-order symmetric system, for barotropic fluids, was found in Special Relativity by K. O. Friedrichs and in General Relativity by A. Rendall[21].

One deduces from (13.1) that

$$\nabla_\alpha u^0 = -\frac{u_i \nabla_\alpha u^i}{u_0}. \tag{13.2}$$

[19] Rendall, A. (1992) *J. Math. Phys.*, **33**, 1047–53.

[20] Anile, A. M. (1987) *Hydrodynamics and Magnetohydrodynamics*, Cambridge University Press.

[21] Rendall, A. (1992) *J. Math. Phys.*, **33**, 1047–53.

The energy equation becomes, using as before the entropy equation

$$\nabla_i u^i - \frac{u_i \nabla_0 u^i}{u_0} + \mu'_p u^\alpha \frac{\partial_\alpha p}{\mu + p} = 0. \tag{13.3}$$

We use the following combination of the equations of motion of indices 0 and i:

$$u^\alpha \nabla_\alpha u^i + (g^{\alpha i} + u^\alpha u^i)\frac{\partial_\alpha p}{\mu + p} - \frac{u^i}{u^0}\left\{u^\alpha \nabla_\alpha u^0 + (g^{\alpha 0} + u^\alpha u^0)\frac{\partial_\alpha p}{\mu + p}\right\} = 0$$

which reduces to, using (13.2),

$$u^\alpha \left(\nabla_\alpha u^i + \frac{u^i u_j}{u^0 u_0}\nabla_\alpha u^j\right) + \left(g^{\alpha i} - \frac{u^i}{u^0} g^{\alpha 0}\right)\frac{\partial_\alpha p}{\mu + p} = 0. \tag{13.4}$$

Multiplying the energy equation (13.3) by $A := (\mu + p)^{-1}$, we find that the principal matrix of the system (13.3), (13.4) for the unknowns p, u^i is, in an arbitrary frame, the following matrix

$$\begin{pmatrix} A^2\mu'_p u^\alpha \partial_\alpha & A(\partial_1 - \frac{u_1}{u_0}\partial_0) & A(\partial_2 - \frac{u_2}{u_0}\partial_0) & A(\partial_3 - \frac{u_3}{u_0}\partial_0) \\ A(\partial^1 - \frac{u^1}{u^0}\partial^0) & u^\alpha(1 + \frac{u^1 u_1}{u^0 u_0})\partial_\alpha & \frac{u^2 u_1}{u^0 u_0}u^\alpha \partial_\alpha & \frac{u^1 u_3}{u^0 u_0}u^\alpha \partial_\alpha \\ A(\partial^2 - \frac{u^2}{u^0}\partial^0) & \frac{u^1 u_2}{u^0 u_0}u^\alpha \partial_\alpha & u^\alpha(1 + \frac{u^2 u_2}{u^0 u_0})\partial_\alpha & \frac{u^2 u_3}{u^0 u_0}u^\alpha \partial_\alpha \\ A(\partial^3 - \frac{u^3}{u^0}\partial^0) & \frac{u^1 u_3}{u^0 u_0}u^\alpha \partial_\alpha & \frac{u^1 u_3}{u^0 u_0}u^\alpha \partial_\alpha & u^\alpha(1 + \frac{u^3 u_3}{u^0 u_0})\partial_\alpha \end{pmatrix} \tag{13.5}$$

Theorem 13.1 *The Euler system written in the form (13.4), (13.5) is symmetrizable. It is hyperbolic with respect to space slices and causal if u is timelike and $\mu'_p \geq 1$.*

Proof In a Cauchy-adapted frame with orthonormal space axis, the metric reads

$$g = -N^2(\theta^0)^2 + \sum_i (\theta^i)^2, \quad \text{with} \quad \theta^0 = dt, \quad \theta^i = \bar{\theta}^i + \lambda^i dt, \quad \bar{\theta}^i = a^i_j dx^j. \tag{13.6}$$

In such a frame it holds that

$$u^0 = -N^{-2}u_0, \quad \partial^0 = -N^{-2}\partial_0, \quad u^i = u_i, \quad \partial^i = \partial_i;$$

the matrix (13.5) takes then the form of a symmetric matrix $\mathcal{M}^0$.

To prove that this symmetrizable system is hyperbolic with respect to space slices it is sufficient to prove (see the section on first-order symmetric hyperbolic systems in Appendix IV) that the matrix $\mathcal{M}^t$ (now symmetric) of the coefficients of $\frac{\partial}{\partial t}$, acting on a vector of R^4 with components P, U_i, is positive definite. This matrix $\mathcal{M}^t$ is identical to the matrix $\mathcal{M}^0$ because, in a Cauchy-adapted frame

$$\partial_0 \equiv \frac{\partial}{\partial t} - \lambda^j \bar{\partial}_i, \quad \partial_i = \bar{\partial}_i.$$

To simplify the computation of the quadratic form associated with the matrix $\mathcal{M}^0$, we choose the first space axis along the projection of u that is we set $u^2 = u^3 = 0$. The matrix $\mathcal{M}^0$ reduces then to

$$\begin{pmatrix} A^2\mu_p' u^0 & -A\frac{u_1}{u_0} & 0 & 0 \\ -A\frac{u_1}{u_0} & u^0(1+\frac{u^1 u_1}{u^0 u_0}) & 0 & 0 \\ 0 & 0 & u^0 & 0 \\ 0 & 0 & 0 & u^0 \end{pmatrix} \tag{13.7}$$

The associated quadratic form, in the variables $P, (U^i)$ is found to be in an arbitrary Cauchy adapted frame (one uses the invariance by space rotations of the frame, and the unitarity of the timelike vector u to obtain this general form)

$$Q(P,V) \equiv u^0 \left\{ A^2\mu_p' P^2 - 2PA\frac{u_i U^i}{u_0 u^0} + U_i U^i + \frac{(u_i U^i)^2}{u_0 u^0} \right\} \tag{13.8}$$

We set

$$W := AP, \quad V := (U^i), \quad v := (u^i), \quad \text{hence} \quad u^0 u_0 = -(1+|v|^2) \tag{13.9}$$

where $|.|$denotes the space norm. The quadratic form Q reads then

$$Q(W,V) \equiv u^0\{(\mu_p' - 1)W^2 + q\} \tag{13.10}$$

where q is the quadratic form

$$q(W,V) := \frac{1}{1+|v|^2}\{(1+|v|^2)W^2 + 2W(v.V) + |V|^2(1+|v|^2) - (v.V)^2\}. \tag{13.11}$$

The Cauchy–Schwarzschild formula implies that

$$q(W,V) \geq \frac{1}{1+|v|^2}\{W^2 + |V+vW|^2\} \geq 0. \tag{13.12}$$

The quadratic form q is positive definite, since the equality $q(W,V) = 0$ implies $W = 0$ hence also $V = 0$. The same is true of Q if $\mu_p' \geq 1$. □

14 Equations in a flow adapted frame

It was long ago stressed by Cattaneo[22] and Ferrarese[23] that the natural physical quantities are time lines and not spacelike submanifolds. Using the same kind of formalism as that given in Chapter 6 for Cauchy-adapted frames, but for Lorentzian frames with time axis tangent to the time lines (**time adapted**

[22] Cattaneo, C. (1959) *Ann. Math. Pura Appl.* **XLVIII**(IV), 361–86.

[23] Ferrarese, G. (1963) *Rendic. Matem.*, **22**, 147–68. See the recent books Ferrarese, G. (2004) *Riferimenti Generalizzati in Relatvità e Applicazioni*, Pitagora, and in English Ferrarese, G. and Bini, D. (2007) *Relativistic Mechanics of Continuous Media*, Springer.

frame) Friedrich[24] has written the perfect fluid equations as a symmetric system for the pressure and the connection coefficients. This system is hyperbolic with respect to a time function under appropriate conditions. We will give the general formulas for the $n+1$ splitting in a time adapted frame, which has other applications, and indicate its use in the perfect fluid case.

14.1 $n+1$ *splitting in a time adapted frame*

14.1.1 *Coframe and metric*

We choose the time axis to be tangent to the time lines, i.e. the cobasis θ is such that θ^i does not contain dx^0. We set

$$\theta^i = a^i_j dx^j, \qquad \theta^0 = U dx^0 + b_i dx^i. \tag{14.1}$$

We will call such a frame a C.F. (**Cattaneo–Ferrarese**) frame. The Pfaff derivatives ∂_α in the C.F. frame are linked to the partial derivatives $\partial/\partial x^\alpha$ by the relations

$$\partial_0 = U^{-1}\frac{\partial}{\partial x^0}, \quad \partial_i = A^j_i\left[\frac{\partial}{\partial x^j} - U^{-1}b_j\frac{\partial}{\partial x^0}\right],$$

with (A^j_i) the matrix inverse of (a^j_i). The structure coefficients of the coframe are found to be:

$$c^0_{0i} = U^{-1}\partial_i U - A^j_i\partial_0 b_j = -c^0_{i0}$$

and, with $f_{[ij]} \equiv f_{ij} - f_{ji}$,

$$c^0_{ij} = UA^h_{[i}\partial_{j]}(N^{-1}b_h),\ c^i_{k0} = A^j_k\partial_0 a^i_j,\ c^i_{hk} = A^j_{[h}\partial_{k]}a^i_j.$$

We choose the frame to be orthonormal, then the metric reads

$$g = -(\theta^0)^2 + \sum_{i=1}^{3}(\theta^i)^2 \tag{14.2}$$

14.1.2 *Splitting of connection*

We deduce from the general formulas that

$$\omega^0_{00} = \omega^0_{i0} = 0,$$

$$Y_i \equiv \omega_{00,i} = \omega^0_{0i} = \omega^i_{00} = -c_{0i,0} = c^0_{0i} = U^{-1}\partial_i U - A^j_i\partial_0 b_j$$

We know that $\omega^j_{0i} = \omega_{oi,j}$ is antisymmetric in i and j. We have

$$\omega_{0i,j} \equiv \frac{1}{2}\{A^h_j\partial_0 a^i_h - A^h_i\partial_0 a^j_i + A^h_{[i}\partial_{j]}b_h\}$$

[24] Friedrich, H. (1998) On the evolution equations for gravitating ideal fluid bodies in General Relativity. *Phys. Rev. D*, **57**, 2317–22. See also Choquet-Bruhat, Y. and York, J. W. (1998) in *Gravitation, Electromagnetism and Geometric Structures* (ed. G. Ferrarese), Pitagora. (These authors use the Riemann tensor, instead of the Weyl tensor used by Friedrich).

Let $e_\alpha \equiv \partial_\alpha$ be the frame dual to θ^α, i.e. such that the vector e_α has components $\delta^\lambda_{(\alpha)}$. Then

$$\nabla_\beta e^\lambda_\alpha = \omega^\lambda_{\beta\alpha};$$

in particular $\omega_{0i,j}$ is the projection on $e_{(j)}$ of the derivative $\nabla_{e_{(0)}} e_{(i)}$ of $e_{(i)}$ in the direction of $e_{(0)}$, $\omega_{0i.j} \equiv (\nabla_{e_{(0)}} e_{(i)}, e_{(j)})$ We choose the spatial frame such that $\omega_{0i,j} = 0$; the formulae are then simplified. Remark that the relation $\omega^0_{0i} = \omega^i_{00}$ written above expresses the vanishing of the derivating along $e_{(0)}$ of the scalar product $(e_{(i)}, e_{(0)})$, i.e. $(\nabla_{e_{(0)}} e_{(i)}, e_{(0)}) = -(\nabla_{e_{(0)}} e_{(0)}, e_{(i)})$. The frame $e_{(\alpha)}$ remains orthonormal under the chosen transport.

The connection coefficient $\omega_{i0,j}$ is the sum of a term symmetric in i and j and an antisymmetric one. We have:

$$X_{ij} \equiv \omega_{i0,j} = \omega^j_{i0} = \frac{1}{2}\{A^h_j \partial_0 a^i_h + A^h_i \partial_0 a^j_i + A^h_{[i} \partial_{j]} b_h\}$$

The antisymmetric term vanishes if the time lines are hypersurface orthogonal ($b_i = 0$).

Remark 14.1 *If the time lines are not hypersurface orthogonal the coefficients c^0_{ij} are different from zero. The coefficients ω^h_{ij} are linear expressions in terms of the first derivatives of the spacetime frame coefficients; they are identical to the connection constructed with the a^j_i at fixed t only if $b_i = 0$.*

14.1.3 *Splitting of curvature*

Using the general formulae we find in the chosen frame

$$R_{0h}{}^i{}_j \equiv \partial_0 \omega^i_{hj} + X^{\cdot\cdot k}_h \omega^i_{kj} + Y^i X^{\cdot\cdot}_{hj} - Y_j X_h{}^i. \tag{14.3}$$

We denote by $\tilde{\nabla}$ the pseudo-covariant derivative constructed from ∂_i and ω^h_{ij} (Cataneo–Ferrarese transversal derivative). We have

$$R_{h0}{}^i{}_0 \equiv \tilde{\nabla}_h Y^i - \partial_0 X^{\cdot\cdot i}_h - X^{\cdot\cdot j}_h X^{\cdot\cdot i}_j, \tag{14.4}$$

$$R_{hk}{}^i{}_j \equiv \bar{R}_{hk}{}^i{}_j + X^{\cdot i}_{\cdot k} X_{jh} - X_{jk} X^i_{\cdot\cdot h}, \tag{14.5}$$

where $\bar{R}_{hk}{}^i{}_j$ denotes the expression formally constructed as a Riemann tensor from the coefficients ω^h_{ij}.

$$R_{khk}{}^0{}_j \equiv \tilde{\nabla}_k X_{hj} - \tilde{\nabla}_h X_{kj} - Y_j (X_{kh} - X_{hk}) \tag{14.6}$$

Remark 14.2 *The symmetry $R_{kh,0j} = R_{0j,kh}$ results from the expression of the connection in terms of the frame coefficients.*

From the splitting of the Riemann tensor we deduce the following identities:

$$R_{00} \equiv \tilde{\nabla}_i Y^i - \partial_0 X^i_i - X^{\cdot\cdot j}_h X^{\cdot\cdot h}_j$$

$$R_{h0} \equiv \tilde{\nabla}_j X^{\cdots j}_h - \tilde{\nabla}_h X^j_j - Y^j (X_{jh} - X_{hj}) \tag{14.7}$$

14.2 *Bianchi equations (case $n = 3$)*

14.2.1 *Bianchi quasi constraints*

The Bianchi identities and their contraction contain, as in a Cauchy-adapted frame, equations which do not contain the derivative ∂_0 of the Riemann tensor. One has the identities

$$\nabla_i R_{jh,\lambda\mu} + \nabla_h R_{ij,\lambda\mu} + \nabla_j R_{hi,\lambda\mu} \equiv 0, \tag{14.8}$$

and the equations, where we replace the Ricci tensor by its value in terms of sources (zero in vacuum),

$$\nabla_\alpha R^\alpha{}_{0,\lambda\mu} = -\nabla_\mu \rho_{\lambda 0} + \nabla_\lambda \rho_{\mu 0}. \tag{14.9}$$

We call these equations Bianchi quasi-constraints.

14.2.2 *Bianchi evolution system*

The remaining Bianchi equations can be written, as in the case of a Cauchy adapted frame, as a FOS (first-order symmetric) system for two pairs of "electric" and "magnetic" 2-tensors. This system cannot be said to be hyperbolic in the usual sense: the principal matrix M^0 is the unit matrix, hence positive definite, but the operators $\partial/\partial t$ appear also in the matrices M^i. We say that the system is a quasi-FOSH system. It is a usual FOSH system with t as a time variable if the matrix of coefficients of $\partial/\partial t$ is positive definite. It can be proved that this is the case if the metric induced on the $t = constant$ submanifolds,

$$\bar{g}_{ij} = \sum_h a_i^h a_j^h - b_i b_j, \tag{14.10}$$

is positive definite, and $U > 0$.

14.2.3 *Quasi-FOSH system for connection and frame*

When the Riemann tensor is known, the identities which express it become equations for the connection. Some of them do not contain the derivative ∂_0 – we call them connection quasi-constraints. Identities linking connection and frame are first-order equations for the frame coefficients. No equation gives the evolution of U. It can be considered as a gauge variable fixing the time parameter.

14.3 *Vacuum case*

In vacuum we give arbitrarily on the spacetime V the scalar U, the length of the tangent vector $\frac{\partial}{dt}$ to the time line, together with the projection Y_i of $\nabla_{e_0} e_0$ on e_i. The quantities f_i^j being chosen zero, the identities previously written give, when the Riemann tensor is known, equations with principal operator the dragging of ω_{hj}^i and X_{ij} along the time lines, and when the connection is known equations for the dragging of the frame coefficients. These equations, together with the Bianchi evolution equations constitute a quasi-FOSH system. This system is a FOSH system with respect to t as long as $\bar{g}_{ij}$ is positive definite.

14.4 *Perfect fluid*

In the presence of fluid sources one can obtain a quasi-FOSH system for the gravitational and fluid variables by taking as time lines the flow lines and proceeding as follows (Friedrich, 1998).

14.4.1 *Fluid equations*

The stress energy tensor of a perfect fluid is

$$T_{\alpha\beta} = (\mu + p)u_\alpha u_\beta + p g_{\alpha\beta}.$$

Hence

$$\rho_{\alpha\beta} = (\mu + p)u_\alpha u_\beta + \frac{1}{2}(\mu - p)g_{\alpha\beta}.$$

For simplicity we suppose that the matter energy density μ is a given function of the pressure p and we denote

$$\mu'_p := \frac{d\mu}{dp}, \quad \mu''_{p^2} := \frac{d^2\mu}{dp^2}.$$

We have seen that the Euler equations of the fluid, which express the conservation law $\nabla_\alpha T^{\alpha\beta} = 0$, are equivalent to the equations

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + (u^\alpha u^\beta + g^{\alpha\beta})\partial_\alpha p = 0, \quad \text{with} \quad u^\alpha u_\alpha = -1$$

and

$$(\mu + p)\nabla_\alpha u^\alpha + u^\alpha \partial_\alpha \mu = 0.$$

In our coframe they read:

$$(\mu + p)Y_i + \partial_i p = 0, \quad Y_i \equiv \omega^i_{00}$$

$$\partial_0 \mu + (\mu + p)X^i_i = 0 \tag{14.11}$$

Using the index of the fluid denoted here F to respect Friedrich's notation

$$F(p) = \int \frac{dp}{\mu(p) + p},$$

we have

$$Y_i = -\partial_i F \tag{14.12}$$

and

$$\mu'_p \partial_0 F + X^i_i = 0. \tag{14.13}$$

The commutation relation between Pfaff derivatives and the definitions give that

$$(\partial_0\partial_i - \partial_i\partial_0)F = c^\alpha_{0i}\partial_\alpha F = Y_i\partial_0 F - X^j_i \partial_j F.$$

Therefore

$$\mu'_p[\partial_0 Y_i + Y_i\partial_0 F + (f^j_i - X^j_i)\partial_j F] - \partial_i \mu'_p \partial_0 F - \partial_i X^h_h = 0 \tag{14.14}$$

The use of previous identities allows us to replace $\partial_\alpha F$ by functions of Y, X, and p. The derivatives $\partial_i \mu'_p$ are functions of Y and p, since

$$\partial_i \mu'_p = \mu''_{p^2} \partial_i p.$$

Following Friedrich we replace in (14.14) $\partial_i X_h^h$ by its expression deduced from the equation

$$R_{i0} \equiv \tilde{\nabla}_h X_i^{\cdot\cdot h} - \tilde{\nabla}_i X_h^h - Y^h (X_{hi} - X_{ih}),$$

and we obtain, changing the names of indices

$$\mu'_p \partial_0 Y_h - \tilde{\nabla}_j X_h^j - Y^h (X_{hi} - X_{ih}) + Y_h \partial_0 F +$$

$$- X_h^j \partial_j F] + \partial_h \mu'_p \partial_0 F = 0 \tag{14.15}$$

In (14.4) we replace $\tilde{\nabla}_h Y_i$ by $\tilde{\nabla}_i Y_h + c^0_{hi} Y_0$ with

$$Y_0 \equiv -\partial_0 F \equiv -(\mu'_p)^{-1} X_i^i,$$

$$c^0_{ih} \equiv \omega^0_{hi} - \omega^0_{ih} \equiv -\omega_{hi,0} + \omega_{ih,0} \equiv X_{ih} - X_{hi} \tag{14.16}$$

and we obtain

$$\partial_0 X_h^i - \tilde{\nabla}^i Y_h + (\mu'_p)^{-1} X_i^i (X_{ih} - X_{hi}) - Y_h Y^i + X_h^{\cdot\cdot j} X_j^{\cdot\cdot i} + X_h^{\cdot\cdot j} = -R_{h0..0}^{\cdot\cdot\cdot\cdot i}. \tag{14.17}$$

The principal operator on the unknowns Y and X in equations (14.15) and (14.17) is diagonal by blocks and symmetric. The h block reads:

$$\begin{pmatrix} \mu'_p \partial_0 & -\partial_1 & -\partial_2 & -\partial_3 \\ -\partial_1 & \partial_0 & 0 & 0 \\ -\partial_2 & 0 & \partial_0 & 0 \\ -\partial_3 & 0 & 0 & \partial_0 \end{pmatrix}$$

If $\mu'_p > 0$ the matrix M^0 is positive definite in the C.F. frame. The system is a quasi-FOSH system for the pairs Y_h, X_h^j.

Remark 14.3 *The characteristic determinant associated with the system is*

$$\left\{ \xi_0^2 \left(\mu'_p \xi_0^2 - \sum_{i=1,2,3} \xi_i^2 \right) \right\}^3$$

The roots of $\mu'_p \xi_0^2 - \sum_{i=1,2,3} \xi_i^2 = 0$ *correspond to sound waves. Their speed is at most 1 (speed of light) if and only if* $\mu'_p \geq 1$.

14.4.2 *Sources of the Bianchi equations*

In our frame the source tensor ρ reduces to

$$\rho_{00} = \frac{1}{2}(\mu + 3p), \quad \rho_{0i} = 0, \quad \rho_{ij} = \frac{1}{2}(\mu - p)\delta_{ij}$$

We have seen that $\partial_\alpha p$ and $\partial_\alpha \mu$ are smooth functions of p, Y and X. The same property holds for $J_{i,0j}$ and $\mathrm{J}_{i,hk}$.

14.5 *Conclusion*

Collecting the results of the previous sections we find the following theorem.

Theorem 14.4 *The Einstein equations with source a perfect fluid give a quasi-FOSH system for the Riemann curvature tensor, the frame and connection coefficients, and the density of matter, when the flow lines are taken as timelines, U and $f^i_j \equiv \omega^i_{0j}$ given arbitrarily.*

Corollary 14.5 *The EEF (Einstein–Euler–Friedrich) system is a FOSH system relatively to $t =$ constant slices as long as the quadratic form*

$$g_{jh} = \sum_{i=1,2,3} a^i_j a^i_h - b_j b_h \tag{14.18}$$

is positive definite, $U > 0$ and $\mu'_p \geq 1$.

15 Charged fluids

15.1 *Equations*

The stress energy tensor of a charged perfect fluid is the sum of the stress energy tensor of the fluid variables and the Maxwell tensor of the electromagnetic field:

$$T_{\alpha\beta} = T^{(fluid)}_{\alpha\beta} + \tau_{\alpha\beta}, \tag{15.1}$$

with

$$T^{(fluid)}_{\alpha\beta} := (\mu + p)u_\alpha u_\beta + p g_{\alpha\beta} \tag{15.2}$$

and

$$\tau_{\alpha\beta} = F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu}. \tag{15.3}$$

The state of the charged fluid is governed by the Einstein equations

$$S_{\alpha\beta} = T_{\alpha\beta}, \tag{15.4}$$

the Maxwell equations, with J the electric current

$$dF = 0 \quad \text{and} \quad \nabla.F = J, \quad \text{i.e.} \quad \nabla_\alpha F^{\alpha\beta} = J^B, \tag{15.5}$$

and the conservation equations which are

$$\nabla_\alpha T^{\alpha\beta} \equiv \nabla_\alpha (T^{\alpha\beta}_{(fluid)} + \tau^{\alpha\beta}) = 0. \tag{15.6}$$

Using previous formulae we see that these conservation equations read

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + g^{\alpha\beta}\partial_\alpha p + u^\beta [\nabla_\alpha (\mu + p)u^\alpha] + J_\lambda F^{\beta\lambda} = 0, \tag{15.7}$$

with, classically, the electric current J sum of a convection and a conduction current:

$$J_\lambda = q u_\lambda + \sigma E_\lambda, \quad E_\lambda = u^\mu F_{\mu\lambda}, \tag{15.8}$$

where q is a scalar function, the electric charge density, σ is the electric conductivity supposed to be constant, and E is the electric field.

The current J satisfies the conservation equation:

$$\nabla_\alpha J^\alpha = 0. \tag{15.9}$$

One deduces from these equations, as for uncharged fluids, by contracted product with u a continuity equation:

$$(\mu + p)\nabla_\alpha u^\alpha + u^\alpha \partial_\alpha \mu - \sigma E^\alpha E_\alpha = 0, \tag{15.10}$$

and equations of motion

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + g^{\alpha\beta}\partial_\alpha p + u^\beta [u^\alpha \partial_\alpha p + \sigma E^\alpha E_\alpha) + J_\lambda F^{\beta\lambda} = 0. \tag{15.11}$$

Modulo initial conditions, as in the vacuum case (see Chapter 6), the Maxwell equations are equivalent to

$$(d\delta + \delta d)F = dJ. \tag{15.12}$$

15.2 *Fluids with zero conductivity*

When the conductivity is zero the Lorentz force reduces to qE, which is orthogonal to the flow vector u, and hence does not furnish any work. One says that the fluid is non dissipative. The Euler and entropy equations are then expected to be a hyperbolic system. This is confirmed by the following theorem.

Theorem 15.1 *The Einstein equations in a wave gauge with sources an electromagnetic field of potential in Lorentz gauge and a charged perfect fluid with equation of state $p = p(\mu, S)$ and with zero conductivity ($\sigma \equiv 0$) are a hyperbolic Leray system for g, A, q, u, and p if $p'_\mu > 0$. This system is causal if $p'_\mu \leq 1$.*

Proof The Einstein equations in a wave gauge with the indicated sources are:

$$R_{\alpha\beta} = R^{(h)}_{\alpha\beta} = F_\alpha{}^\lambda F_{\beta\lambda} - \frac{1}{4} g_{\alpha\beta} F^{\lambda\mu} F_{\lambda\mu} + (\mu + p)u_\alpha u_\beta + \frac{1}{2} g_{\alpha\beta}(\mu - p) \tag{15.13}$$

The Maxwell equations with electromagnetic potential A in Lorentz gauge read as wave equations for A:

$$\nabla_\alpha \partial^\alpha A^\beta - R^\beta{}_\lambda A^\lambda = J^\beta = q u^\beta, \tag{15.14}$$

In the case where $\sigma = 0$ Equation (15.10) reduces to the continuity equation of an uncharged fluid

$$\nabla_\alpha [(\mu + p)u^\alpha] - u^\alpha \partial_\alpha p = 0, \tag{15.15}$$

as could be foreseen from the fact that the Lorentz force is orthogonal to u. The particle number conservation equation is, as in an uncharged fluid, equivalent to the entropy conservation along the flow lines

$$u^\alpha \partial_\alpha S = 0. \tag{15.16}$$

In the equations of motion, the Lorentz electric force qE appears. These equations read

$$(\mu + p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\beta u^\alpha)\partial_\alpha p - qE^\beta = 0. \tag{15.17}$$

The conservation equation of the electric current reduces to a conservation of charge equation:

$$\nabla_\alpha (qu^\alpha) \equiv u^\alpha \partial_\alpha q + q\nabla_\alpha u^\alpha = 0. \tag{15.18}$$

We can combine this equation with the particle number conservation equation to show that q/r is constant along the flow lines:

$$u^\alpha \partial_\alpha \frac{q}{r}. \tag{15.19}$$

To keep as independent unknowns μ and S, we introduce the function $\tilde{f}(\mu, S)$ defined by:

$$\tilde{f}(\mu, S) =: exp \int_{\mu_0}^{\mu} \frac{dm}{m + p(m, S)} \tag{15.20}$$

Then

$$\partial_\alpha \tilde{f} = \tilde{f} \left\{ \frac{\partial_\alpha \mu}{\mu + p} - \int_{\mu_0}^{\mu} \frac{p'_S(m, S)}{[m + p(m, S)]^2} dm \right\} \partial_\alpha S. \tag{15.21}$$

Modulo the conservation of entropy equation, the conservation of charge reads

$$u^\alpha \partial_\alpha \tilde{q} = 0, \quad \text{with } \tilde{q} := \tilde{f}^{-1} q. \tag{15.22}$$

For each pair μ, S, and when the equation of state $p(\mu, S)$ is known, $\tilde{q}$ is a known function of q and conversely. We take μ, S, $\tilde{q}$ as independent unknowns.

Choose the Leray–Volevic indices to be zero for all equations, two for the unknowns g and A and one for the other unknowns; then the characteristic matrix of the system of equations for g, A, $\tilde{q}$, S, μ, u is the matrix diagonal by blocks

$$\mathcal{M}_{\sigma,\sigma=0} = \begin{pmatrix} g^{\alpha\beta} X_\alpha X_\beta I_{14} & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & \mathcal{M} \end{pmatrix} \tag{15.23}$$

where I_{14} is the unit 14×14 matrix, $a := u^\alpha X_\alpha$ and $\mathcal{M}$ is the matrix given in Section 8.2.

It is straightforward, using previous results on $\mathcal{M}$, to check that the system is a Leray hyperbolic system, causal if $\mu'_P > 0$. Apart from the addition of electromagnetic wave fronts, the wave fronts are the same as for uncharged fluids.

An existence theorem for the solution of the full Einstein–Maxwell charged fluid, for initial data satisfying the constraints, can be deduced by methods analogous to the ones used before. □

16 Fluids with finite conductivity

When the conductivity σ is finite and non zero the Lorentz force is no longer orthogonal to u: it works under the fluid flow. The fluid may be called dissipative and its properties are expected to be different from those of non-dissipative fluids. Indeed, experiments show, for example, that no shock wave propagates in charged mercury, a liquid with non-zero conductivity. This fact leads to the thought that for such a fluid the Cauchy problem is not well posed in spaces of functions with a finite number of derivatives, namely in Sobolev spaces. Fluids with finite, non-zero conductivity are a physical example of the Leray–Ohya theory of non strictly hyperbolic systems[25].

When the conductivity is not zero, the entropy is not conserved along flow lines: we deduce from the conservation of energy and the conservation of matter that

$$u^\alpha \partial_\alpha S = \sigma E^\alpha E_\alpha \geq 0. \tag{16.1}$$

Hence the entropy increases[26] along the flow lines.

The conservation of the electric current J,

$$\nabla_\alpha (q u^\alpha + \sigma u_\lambda F^{\alpha\lambda}) = 0, \tag{16.2}$$

contains derivatives of the electromagnetic field F, hence second derivatives of the potential A. We give to the equation above the Leray–Volevic index 0 and to q the index 1. The characteristic matrix is no more diagonal by blocks. It is of the form:

$$\mathcal{M}_\sigma = \begin{pmatrix} g^{\alpha\beta} X_\alpha X_\beta I_{10} & 0 & 0 & 0 & 0 \\ 0 & g^{\alpha\beta} X_\alpha X_\beta I_4 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 \\ 0 & \sigma P_A(X) & 0 & a & P_u(X) \\ 0 & 0 & 0 & 0 & \mathcal{M} \end{pmatrix} \tag{16.3}$$

where $P_A(X)$ and $P_u(X)$ are matrices with four columns and one row with elements respectively quadratic and linear in X, deduced from the second derivatives of the $A's$ and the first derivatives of the $u's$ in the equations[27].

The characteristic determinant is the same as in the case of zero conductivity, but the quasi-diagonalisation theorem does not apply. The system is causal if $p'_\mu \leq 1$, but only hyperbolic non-strict.

[25] Choquet-Bruhat, Y. (1965) *C. R. Acad. Sci. Paris*, **261**, 354–6.

[26] Recall that E is a spacelike vector.

[27] We could replace q by another function $\tilde{q}$, as we did in the case of zero conductivity. It will not transform the equation of conservation of the electric current into a differential equation for $\tilde{q}$ along the flow lines, but it will give a factor also in front of P_u.

Theorem 16.1 *The reduced Einstein–Maxwell–Euler entropy system of a charged perfect fluid with finite, non-zero, conductivity is hyperbolic (and causal) but only in the Leray–Ohya sense (hyperbolic non-strict)*[28].

17 Magnetohydrodynamics

In the case where the electric conductivity is so large that it can be considered as infinite, the electric field becomes negligible. The case of a zero electric field E is called magnetohydrodynamics[29]. It plays a fundamental role in plasma physics.

17.1 *Equations*

The equations of magnetohydrodynamics in a spacetime of General Relativity have been first written, and the wave fronts computed, in Choquet-Bruhat (1960)[30].

In three space dimensions, for a fluid with **infinite conductivity**, $\sigma = \infty$, the electromagnetic field reduces to the magnetic vector H. The second Maxwell equation $\delta F = J$ is replaced by

$$E^\alpha := u_\beta F^{\beta\alpha} = 0. \tag{17.1}$$

From the definition of the vector H (orthogonal to u)

$$H^\alpha = u_\beta (F^*)^{\beta\alpha} \tag{17.2}$$

one deduces

$$(F^*)^{\beta\alpha} = H^\alpha u^\beta - H^\beta u^\alpha. \tag{17.3}$$

A straightforward calculation gives the following expression for the Maxwell tensor

$$\tau_{\alpha\beta} \equiv \left(u_\alpha u_\beta - \frac{1}{2} g_{\alpha\beta} \right) |H|^2 - H_\alpha H_\beta \tag{17.4}$$

The first Maxwell equation $dF = 0$ becomes

$$\nabla_\alpha (F^*)^{\beta\alpha} \equiv \nabla_\alpha (H^\alpha u^\beta - H^\beta u^\alpha) = 0. \tag{17.5}$$

Modulo these equations, the divergence of τ is found to be

$$\nabla_\alpha \tau^{\alpha\beta} = \left(u^\alpha u^\beta - \frac{1}{2} g^{\alpha\beta} \right) \partial_\alpha |H|^2 + |H|^2 \nabla_\alpha (u^\alpha u^\beta) - \nabla_\alpha (H^\alpha H^\beta) \tag{17.6}$$

[28] O. Friedrichs has obtained for a charged perfect fluid with non-zero finite conductivity a symmetric hyperbolic system, but by modifying the equations in a way not physically justified.

[29] Equations for magnetohydrodynamics in Special Relativity, keeping terms neglected in classical magnetohydrodynamics, were obtained by Hoffman, F. and Teller, E. (1950), *Phys. Rev. Ser. 2*, **80**, 692–703.

[30] Choquet-Bruhat, Y. (1960) *Astron. Acta*, **VI**(6), 354–65.

Straightforward computation shows that it is a vector orthogonal to both u and H.

The Lorentz force $\nabla_\alpha \tau^{\alpha\beta}$ being orthogonal to u, the continuity equation for a fluid of infinite conductivity is the same as for an uncharged fluid, namely

$$\nabla_\alpha[(\mu+p)u^\alpha] - u^\alpha \partial_\alpha p = 0, \tag{17.7}$$

The equations of motion read

$$(\mu+p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\beta u^\alpha)\partial_\alpha p + \nabla_\alpha \tau^{\alpha\beta} = 0. \tag{17.8}$$

17.2 *Wave fronts*

A straightforward computation shows that the characteristic polynomial of the first order system (17.5), (17.7), (17.8) is

$$[(\mu+p)a^2 - b^2]^2 D, \quad \text{with} \quad a := u^\alpha X_\alpha, \quad b := H^\alpha X_\alpha, \tag{17.9}$$

and with

$$D \equiv (\mu'_p - 1)(\mu+p)a^4 + [(\mu+p+|H|^2\mu'_p)a^2 - b^2]X^\alpha X_\alpha. \tag{17.10}$$

The cone

$$(\mu+p)a^2 - b^2 = 0 \tag{17.11}$$

is called the **Alfvén cone**. It is composed of two hyperplanes:

$$[(\mu+p)^{\frac{1}{2}} u^\alpha \pm H^\alpha]X_\alpha = 0, \tag{17.12}$$

The corresponding wave fronts, tangent to the dual of a Alfvén plane, are called the **Alfvén waves**.

In the proper frame of the fluid the normals ν to the Alfvén waves satisfy the equations

$$(\mu+p)^{\frac{1}{2}} \nu_0 \pm H^i \nu_i = 0, \tag{17.13}$$

The propagation speed of the Alfvén waves are computed using the general formula

$$|V|^2 = \frac{(u^\alpha \nu_\alpha)^2}{(u^\alpha u^\beta + g^{\alpha\beta})\nu_\alpha \nu_\beta}. \tag{17.14}$$

One finds that

$$|V_{Alf}| = (\mu+p)^{-\frac{1}{2}} |H_{\bar{\nu}}| \tag{17.15}$$

where H_ν is the scalar product of H (a space vector) with the projection on space (normed to 1) of the normal ν to the wave front.

The wave fronts whose normals lie on the magnetoacoustic cone $D = 0$ are called the **magnetoacoustic** wave fronts. The equation of the magnetoacoustic

cone is a polynomial of order four in ν, the normal to a wave front, which reads in a proper rest frame of the fluid

$$\mu'_p(\mu + p + H^2)(\tilde{\nu}_0)^4 - \tilde{\nu}_0^2\{(\mu + p + |H|^2\mu'_p + H_{\tilde{\nu}}^2\} + {H_{\tilde{\nu}}}^2 = 0 \tag{17.16}$$

where we have denoted

$$\tilde{\nu} = \frac{\nu}{\sum_i [\nu_i^2]^{1/2}} \quad \text{and} \quad H_{\tilde{\nu}} = H_i\tilde{\nu}^i. \tag{17.17}$$

We deduce from this formula, solved for the unknown $\tilde{\nu}_0^2$, the two magnetoacoustic wave speeds. Both are well defined because $|H_{\tilde{\nu}}| \leq |H|$, and less than 1, the speed of light, if $\mu'_p \geq 1$. They are

$$|V|^2_{MA} = \frac{(\mu + p + |H|^2\mu'_p + H_{\tilde{\nu}}^2 \pm \{[(\mu + p + |H|^2\mu'_p + H_{\tilde{\nu}}^2]^2 - 4H_{\tilde{\nu}}^2\mu'_p(\mu + p + H^2)\}^{\frac{1}{2}}}{2\mu'_p(\mu + p + H^2)}. \tag{17.18}$$

The rapid wave speed corresponds to the plus sign, the slow wave speed to the minus sign.

The study of the characteristic determinant shows[31] that the system is a hyperbolic system of the Leray–Ohya type. The fact that Alfvén wave fronts are tangent to the magnetosonic wave fronts does not permit us to conclude that it has (strict) Leray hyperbolicity. However, the system has been proved to be symmetrizable hyperbolic, both by the general method of Lax-Boillat and Ruggeri, and directly, at least in special Relativity, by O. Friedrichs. This strict hyperbolicity is in agreement with the existence of shock waves[32] and of high frequency waves[33] propagating by first-order differential equations.

18 Yang–Mills fluids

Although Yang–Mills charges are not manifest at ordinary scales, plasmas of quarks and gluons exist in extreme circumstances. Their properties have analogies with, but also differences from, electromagnetic plasmas. The equations look formally the same. The Maxwell stress energy tensor is replaced by the Yang–Mills stress energy tensor and the electric current by the Yang–Mills current

$$J^\alpha = \gamma u^\alpha + \sigma F^{\beta\alpha}u_\beta,$$

where γ is a function on spacetime taking its values in the same Lie algebra as the Yang–Mills field F and σ is a number, the fluid conductivity.

[31] See Choquet-Bruhat, Y. (1966) *Commun. Math. Phys.*, **3**, 334–57.

[32] General relativistic equations given in Choquet-Bruhat, Y. (1960) *Astron. Acta*, **VI**, 354–65. Studied in depth in Lichnerowicz (1967).

[33] Anile, M. and Greco, A. (1979) *Nuovo Cimento B*, **55**(2), 307–24.

The mathematical properties of **Yang–Mills fluids** are[34] quite similar to those of electrically charged fluids, at least locally, in the case of finite conductivity.

In the case of **infinite conductivity** a remarkable property occurs: the wave fronts do not split into the analogues of Alfvén waves and acoustic waves, but are at each point tangent to an undecomposable sixth-order cone[35].

19 Dissipative fluids

There have been many discussions and proposals for the writing of equations for relativistic viscous fluids, and for a relativistic heat equation. Most proposals were motivated by the desire to obtain a causal theory; that is, partial differential equations of hyperbolic type with a causal domain of dependence, while the usual Fourier law of heat transfer, as well as viscous fluids Navier–Stokes equations of Newtonian mechanics are of parabolic type, corresponding to an infinite propagation speed. We will see in Chapter 10 some results for the approximation beyond perfect fluids obtained in the framework of Relativity.

19.1 *Viscous fluids*

The strong equivalence principle would give us an expression for the stress energy tensor of a viscous fluid in General Relativity if we knew it in Special Relativity, but this is not the case. There is no general consensus about such a tensor, everyone being influenced in his choice by his own background. The difference between various authors begins with the physical definition of the flow vector[36]. We will stick to the simplest definition, considering r as the particle number density and u the flow vector. These quantities satisfy the conservation law

$$\nabla_\alpha (r u^\alpha) = 0. \tag{19.1}$$

It was remarked previously that the components T^{00}, T^{0i}, T^{ij} of the stress energy tensor T in a proper frame represented respectively the energy, momentum, and stresses of the fluid with respect to this frame. For perfect fluids, the expressions for these quantities were taken from their classical analogues, which result themselves from first principles.

19.1.1 *Navier–Stokes equations*

In non-relativistic mechanics, the stress tensor of a viscous fluid is obtained by considering it as a linearized perturbation σ of the perfect fluid stress tensor, the

[34] Choquet-Bruhat, Y. (1992) *J. Math. Phys.*, **33**(5), 1782–5 and (1996) Lecture Notes in Physics 460, pp. 3–31.

[35] Choquet-Bruhat, Y. (1992) *C. R. Acad. Sci Paris*, serie 1 **318**, 775–82, and in Rionero, S. and Ruggeri, T. (1992) (eds.) *Waves and Stability in Continuous Media*, pp. 54–69, World Scientific, Singapore.

[36] Landau and Lifshitz link u with some density of energy, choosing u such that the stress tensor is orthogonal to u.

perturbation being due to the deformation tensor of the flow lines. In inertial coordinates of absolute space this deformation tensor is

$$D_{ij} := \frac{\partial v_i}{\partial x^j} + \frac{\partial v_j}{\partial x^i}. \tag{19.2}$$

One writes the **Navier–Stokes** equations for viscous fluids by introducing the expansion and shear of the flow lines (in Euclidean space E^3). The rotational component is discarded, argued as corresponding to rigid motions with do not generate stress. The Newtonian stress tensor of a viscous fluid is then the sum of the stress tensor $p\delta_{ij}$ of a perfect fluid and a symmetric 2-tensor given by

$$\sigma_{ij} := \lambda\delta_{ij}\frac{\partial v^k}{\partial x^k} + \nu\left(D_{ij} - \frac{2}{3}\delta_{ij}\frac{\partial v^k}{\partial x^k}\right). \tag{19.3}$$

The scalar coefficients, in general considered as constants, are called the bulk viscosity and the shear viscosity.

The extension of the tensor σ to Relativity is ambiguous, due on the one hand to the absence of absolute space and on the other hand to the equivalence of mass with energy, hence with the work of the friction forces due to viscosity.

The following tensor has been proposed as a simple generalization of the classical one:

$$T_{\alpha\beta} := (\mu + p)u_\alpha u_\beta + pg_{\alpha\beta} + \sigma_{\alpha\beta} \tag{19.4}$$

with σ the viscosity stress tensor, linear in the first derivatives of u and orthogonal to u

$$\sigma_{\alpha\beta} = \lambda\pi_{\alpha\beta}\nabla_\rho u^\rho + \nu\pi^\rho_\alpha\pi^\sigma_\beta(\nabla_\rho u_\sigma + \nabla_\rho u_\sigma) \tag{19.5}$$

where π is the projection tensor

$$\pi_{\alpha\beta} := g_{\alpha\beta} + u_\alpha u_\beta. \tag{19.6}$$

Some authors add to T a "momentum term"

$$u_\alpha q_\beta + u_\beta q_\alpha, \tag{19.7}$$

supposed to represent a heat flow resulting from friction. The problem is then to write an equation to determine q and justify the symmetrization in (19.7).

The equations $\nabla_\alpha T^{\alpha\beta}$ deduced from (19.4) and (19.5) are rather complicated. They have been studied in detail by G. Pichon in the case where $\nabla_\alpha u^\alpha = 0$, called an incompressible fluid. He has shown that for large enough $\mu + p$ they are of parabolic type, corresponding therefore to infinite propagation speed.

19.1.2 *A hyperbolic relativistic Navier–Stokes system*

We have seen that the vector which leads to the generalization of classical properties of irrotational motion, for perfect fluids, is the dynamic velocity $C := hu$.

We define for this vector C the shear, the vorticity tensor, and the expansion of the congruence of its trajectories (the flow lines of the fluid) by the usual decomposition (see Chapter 1) in the case $n = 3$,

$$\nabla_\alpha C_\beta = \Omega_{\alpha\beta} + \Sigma_{\alpha\beta} + \frac{1}{3}\Theta g_{\alpha\beta},$$

with Ω the vorticity tensor:

$$\Omega_{\alpha\beta} := \frac{1}{2}\{\nabla_\rho C_\sigma - \nabla_\rho C_\sigma\}.$$

In the case of a perfect fluid Ω is orthogonal to C (i.e. to u). The scalar Θ is the expansion of the congruence defined by the vector field C:

$$\Theta := \nabla_\alpha C^\alpha.$$

We have seen that Θ is zero if and only if the fluid is incompressible (in the relativist sense). The symmetric tensor Σ is the shear, which has a zero trace:

$$\Sigma_{\alpha\beta} := \frac{1}{2}\{\nabla_\alpha C_\beta + \nabla_\beta C_\alpha\} - \frac{1}{4} g_{\alpha\beta}\nabla_\lambda C^\lambda.$$

We propose to take as stress energy tensor of a viscous fluid the tensor

$$T_{\alpha\beta} = (\mu + p)u_\alpha u_\beta + g_{\alpha\beta}p + \tilde{T}_{\alpha\beta}, \tag{19.8}$$

where $\tilde{T}_{\alpha\beta}$ is the part of T due to viscosity, a perturbation of the perfect fluid stress energy tensor, linear in ∇C, of the form

$$\tilde{T}_{\alpha\beta} := \lambda\Theta g_{\alpha\beta} + \nu\Sigma_{\alpha\beta}. \tag{19.9}$$

with λ and ν viscosity coefficients depending on the considered fluid. The conservation laws $\nabla_\alpha T^{\alpha\beta} = 0$ are then second-order equations for the dynamical velocity C with principal part

$$\frac{\nu}{2}\{\nabla_\alpha\nabla^\alpha C^\beta + \nabla_\alpha\nabla^\beta C^\alpha\} + \left(\lambda - \frac{1}{3}\nu\right)\nabla^\beta\nabla_\alpha C^\alpha.$$

The 4×4 characteristic matrix has elements

$$\frac{\nu}{2}X_\alpha X^\alpha C^\beta + aX_\alpha X^\beta C^\alpha \qquad a := \frac{\nu}{6} + \lambda.$$

To compute the characteristic determinant we choose a frame where $X^2 = X^3$. The matrix then reads

$$\begin{pmatrix} \frac{\nu}{2}X^\alpha X_\alpha + \alpha X_0 X^0 & aX_1 X^0 & 0 & 0 \\ aX_0 X^1 & \frac{\nu}{2}X^\alpha X_\alpha + \alpha X_1 X^1 & 0 & 0 \\ 0 & 0 & \frac{\nu}{2}X^\alpha X_\alpha & 0 \\ 0 & 0 & 0 & \frac{\nu}{2}X^\alpha X_\alpha \end{pmatrix}. \tag{19.10}$$

The characteristic determinant is found to be, returning to a general frame,

$$\frac{\nu}{2}\left(\frac{\nu}{2} + a\right)(X^\alpha X_\alpha)^4. \tag{19.11}$$

The proposed system is Leray–Ohya hyperbolic (Gevrey class of index 2) and causal, a satisfactory property for a relativistic dissipative theory.

It can be conjectured that a solution of the proposed relativistic Navier–Stokes equations converges to a solution of the perfect fluid equations when λ and ν tend to zero.

19.2 *The heat equation*

We personally think that heat transfer is a collective effect; the Fourier equation corresponds in fact to some asymptotic steady state.[37] It is not surprising that it does not translate readily into relativistic causal equations. The classical heat equation appears after the study of Brownian motion, but unfortunately the extension to Relativity of a study of Brownian motion is still in its infancy due to several difficulties, in particular the lack of absolute time.

The Muller–Ruggeri extended thermodynamics offers a way to treat dissipative phenomena in a relativistic context (see Chapter 10).

[37] A singular perturbation with higher derivative has been proposed by Vernotte to give an account of the few moments occurring before the steady state is attained. Vernotte's idea has been extended to General Relativity by Cattaneo.

X

RELATIVISTIC KINETIC THEORY

1 Introduction

We deduced in Chapter 2 the stress energy tensor of a perfect fluid from the dynamical law of Special Relativity. The equivalence principle (see Chapter 3) leads to the same stress energy tensor for a perfect fluid in General Relativity, and to the relativistic Euler equations by the vanishing of the covariant divergence of this stress energy tensor in the spacetime metric. We have seen (Chapter 9) that the relativistic theory of perfect fluids is satisfactory in many respects, physically[1] as well as mathematically. Global mathematical problems remain mostly open, but it is also the case for classical (i.e. non-relativistic) fluids in spite of recent progress. A new book by Christodolou[2] treats four-dimensional relativistic perfect fluids in depth up to shock formation, but the structure of four-dimensional fluids after shock formation is still difficult to handle mathematically.

We also pointed out in Chapter 9 that there are no compelling macroscopic considerations leading to satisfactory relativistic equations for dissipative fluids, even in Special Relativity. The classical Navier–Stokes equations lead, like the Fourier law of heat transfer, to equations of parabolic type, hence to an infinite propagation speed of signals, which is incompatible with relativistic causality. The formal generalizations of these equations in view of obtaining hyperbolic equations lack general justification and there is no consensus about them.

An *a posteriori* justification of the non-relativistic Euler and Navier–Stokes macroscopic equations is to deduce them from the motion of the fluid particles at the microscopic scale and the statistical hypothesis of the classical kinetic theory. It presents no conceptual difficulty to extend the setting of kinetic theory to Special as well as to General Relativity. It is straightforward to construct an energy momentum vector and a stress energy tensor for kinetic matter and to couple the last one with the Einstein equations. The collective motion of collisionless particles is naturally modelled and leads to conservation equations for the stress energy tensor. A case of particular interest in General Relativity is when the "particles" are stars, galaxies, or clusters of galaxies. It is then

[1] Although there is some problem for the choice of an equation of state for the fluid, particularly when the whole cosmos is concerned.

[2] Christodoulou, D. (2007) *The Formation of Shocks in Relativistic Fluids*, EMS, Zurich.

appropriate to take the charges to be zero and the masses to vary between two positive numbers.

When the particles undergo collisions one can write a Boltzmann equation, and a coherent Einstein–Boltzmann system, for an appropriate choice of the collision cross-section. The problem is the choice of a physically reasonable cross-section, but this problem also arises in the non-relativistic case. The Boltzmann equation is particularly interesting for the study of plasmas of elementary particles with a finite number of distinct proper masses and charges, and, possibly, obtaining equations for dissipative relativistic fluids.

In Sections 2–6 we establish the basic equations of the relativistic kinetic theory[3]; we prove local existence theorems for Einstein equations coupled with kinetic matter[4]. In Section 7 we give thermodynamic properties linked with the Boltzmann equation, we prove the H-theorem, and we indicate how perturbation around a Maxwell–Jütner equilibrium distribution gives possible equations for dissipative fluids. The last section, contributed by T. Ruggeri, indicates how the theory of extended thermodynamics circumvents the difficulty of generalizing to Relativity the dissipative fluids equations.

2 Distribution function

2.1 *Definition*

In kinetic theory matter is composed of a collection of particles whose size is negligible at the considered scale: rarefied gases in the laboratory, galaxies or even clusters of galaxies at the cosmological scale. The number of particles is so great and their motion so chaotic that it is impossible to observe their individual motion. It is assumed that the state of the matter in a spacetime (V, g) is represented[5] by a "one-particle distribution function". This distribution function f is interpreted as the density of particles at a point $x \in V$ which have a momentum $p \in T_xV$, the tangent space to V at x; the vector p is future timelike or null. We state a definition.

Definition 2.1 *A* ***distribution function*** f *is a positive scalar function on the so-called* ***phase space*** $\mathcal{P}_V$ *subbundle of the tangent bundle* $T(V)$ *to the spacetime* V*,*

$$f : \mathcal{P}_V \to R \quad by \quad (x,p) \mapsto f(x,p), \quad with \quad x \in V, \quad p \in \mathcal{P}_{\mathbf{x}} \subset T_xV. \tag{2.1}$$

the fibre $\mathcal{P}_{\mathbf{x}}$ *at* x *is such that* $g_x(p,p) \leq 0$ *and in a time-oriented frame* $p^0 \geq 0$*.*

[3] Basic references are the two papers by Marle, C. M. (1969) *Ann. Inst. Poincaré* **2**, 67–126 and 127–94. These papers contain references to earlier works.

[4] A brief survey of what is known at present about global existence of solutions of the Einstein equations with kinetic source is given in Section 5.3.2.

[5] A mathematical justification of the onset of chaos in relativistic dynamics is a largely open problem.

If the particles have all the same mass m, then the phase space denoted $\mathcal{P}_{m,V}$ has for fibre $\mathcal{P}_{m,x}$ the **mass hyperboloid**

$$\mathcal{P}_{\mathbf{m},\mathbf{x}} \equiv T_x V \cap \{g(p,p) = -m^2, p^0 > 0\}. \tag{2.2}$$

Remark 2.2 *If the particles have a positive rest mass, then it holds that*

$$p^0 > 0 \tag{2.3}$$

($p^0 \geq m$ if the frame is orthonormal).

If the particles have a zero rest mass, then p^0 can vanish only if p itself vanishes. The value $p = 0$ is from the mathematics point of view, a singular point of the vector field X. On the other hand, relativistic physics says that particles with zero mass move with the speed of light, the component p^0 in an orthonormal frame is the energy with respect to an observer at rest in this frame, quantum theory tells us that it does not vanish.

2.2 *Interpretation*

The physical meaning of the distribution function is that it gives a mean[6] "presence number" density of particles in phase space. More precisely we denote θ the volume $2(n+1)$-form on TV, i.e. with ω_g and ω_p respectively the volume forms on V and T_xV:

$$\theta = \omega_g \wedge \omega_p. \tag{2.4}$$

In local coordinates ω_g and ω_p are given by:

$$\omega_g = (\det g)^{\frac{1}{2}} dx^0 \wedge dx^1 \wedge \cdots \wedge dx^n, \tag{2.5}$$

and

$$\omega_p := (\det g)^{\frac{1}{2}} dp^0 \wedge dp^1 \cdots \wedge dp^n. \tag{2.6}$$

In the case of particles of a given mass m the volume form $\omega_{m,p}$ in $\mathcal{P}_{m,x}$ is (Leray form) such that

$$d\left[\frac{1}{2}(g_{\alpha\beta} - m^2)\right] \wedge \omega_{m,p} = \omega_p, \tag{2.7}$$

taking the p^i as local coordinates on $\mathcal{P}_{m,x}$ (then p^0 is a function of x and p^i) gives

$$\omega_{m,p} = \frac{(\det g)^{\frac{1}{2}} d^n p}{p_0}, \quad d^n p := dp^1 \wedge \cdots \wedge dp^n. \tag{2.8}$$

The volume form in $\mathcal{P}_m$ is $\theta_m = \omega_g \wedge \omega_{m,p}$.

Throughout this chapter we assume that the spacetime V is an oriented manifold, the forms $f\theta$ and $f\theta_m$, induce respectively the measure elements for the mean presence number of particles in domains of respectively $\mathcal{P}_V$ and $\mathcal{P}_m$, which are oriented by the orientation of V, all fibres are positively oriented.

[6] In the sense of Gibbs ensemble.

2.3 *Moments of the distribution function*

The moments of f are functions or tensors on V obtained by integration on the fibers of the phase space $\mathcal{P}_V$ of products of f by tensor products of p with itself. We denote by μ_g and μ_p the volume elements[7] associated to the volume forms ω_g and ω_p.

2.3.1 *Moment of order zero*

It is by definition the integral on the fibre $\mathcal{P}_x$ of the distribution function:

$$r_0(x) := \int_{\mathcal{P}_x} f\ \omega_p = \int_{\mathcal{P}_x} f\ \mu_p \tag{2.9}$$

It is a "density of presence" in spacetime. An analogous definition holds for particles of given mass m.

Remark 2.3 *Given a distribution function f and a $2n+1$-dimensional submanifold Σ of the phase space $\mathcal{P} \subset TV$ the mean number N_Σ of trajectories crossing Σ is*

$$N_\Sigma = \int_\Sigma f\ i_X\theta, \tag{2.10}$$

where $X = (p, Q)$ is the vector at a point $(x, p) \in \Sigma$ tangent to the trajectory in $\mathcal{P}_V$ of particles crossing Σ at this point. We choose for Σ a subbundle $\mathcal{P}_S$ of $\mathcal{P}_V$ with base S an n-dimensional submanifold S of the basis V of $\mathcal{P}_V$. Its fibres are $\mathcal{P}_x$, $x \in S$. We recall the property of the interior product (CB-DM1 IV A 4) when ω_g is a form of degree $n+1$:

$$i_X\theta \equiv i_X(\omega_g \wedge \omega_p) = i_X\omega_g \wedge \omega_p + (-1)^{n+1}\omega_g \wedge i_X\omega_p. \tag{2.11}$$

The bundle Σ is locally a product $S \times R^{n+1}$. Since S is n-dimensional and η is a $n+1$ form it induces a zero form on S. The form induced on Σ by $i_X\theta$ reduces therefore to

$$(i_X\theta)_\Sigma = (i_X\omega_g)_S \wedge \omega_p = p^\alpha n_\alpha \bar{\omega}_{g,S} \wedge \omega_p \tag{2.12}$$

where n_α is the unit normal to S and $\bar{\omega}_{g,S}$ is the volume form of S in the metric induced by the Lorentzian spacetime metric ω_g. We deduce from the formulas (2.10) and (2.11) that the quantity

$$r_S(x) = \int_{\mathcal{P}_x} f\ p^\alpha n_\alpha \omega_p \tag{2.13}$$

is to be interpreted as the mean density of trajectories crossing S at the point x.

[7] Volume forms are anticommutative, volume elements are commutative (Fubini theorem). For integration of exterior forms on oriented manifolds see for instance CB-DM1 IV B.

2.3.2 *First and second moments*

The **first moment** of f is a vector field on V defined by

$$P^\alpha(x) := \int_{\mathcal{P}_x} p^\alpha f(x,p)\mu_p. \tag{2.14}$$

If the spacetime is time-oriented and the particles travel towards the future, i.e. if p^α is causal (timelike or null) with $p^0 > 0$, the fibre $\mathcal{P}_x$ is included in the subset $p^0 > 0$. The vector fp^α is then also causal and future directed since $f \geq 0$, and the same is true of P^α.

Out of the first moment one extracts a scalar $r \geq 0$ interpreted as the square of a specific proper mass density given by:

$$r^2 := -P^\alpha P_\alpha. \tag{2.15}$$

If P is time-like, then $r > 0$. One deduces from the first moment a unit vector u interpreted as the macroscopic flow velocity given by

$$u^\alpha := r^{-1} P^\alpha. \tag{2.16}$$

A sufficient condition for P to be timelike is that the particles have a positive mass since then all p, hence also P are timelike.

The **second moment** of the distribution function f is the symmetric 2-tensor on spacetime given by

$$T^{\alpha\beta}(x) := \int_{\mathcal{P}_x} f(x,p) p^\alpha p^\beta \mu_p. \tag{2.17}$$

It is interpreted as the stress energy tensor of the distribution f.

Theorem 2.4 *If a distribution function f depends on p only through a scalar product $V_\alpha p^\alpha$ with V^α a given timelike vector on spacetime, then the first and second moments of f read respectively as the particle number–momentum vector and the stress energy tensor of a perfect fluid. The unit flow velocity is collinear to V.*

Proof Assume that

$$f(x,p) \equiv F(x, V_\alpha(x) p^\alpha), \tag{2.18}$$

the first moment of f is then

$$P^\alpha(x) \equiv \int_{\mathcal{P}_x} p^\alpha F(x, V_\alpha(x) p^\alpha) \mu_p. \tag{2.19}$$

Take at x an orthonormal Lorentzian frame with time axis U collinear to V. In this frame $V_\alpha(x) = -\lambda(x)\delta^0_\alpha$, with $\lambda := (-V_\alpha V^\alpha)^{\frac{1}{2}}$, and $\mu_p = dp^0 dp^1 \dots dp^n$ hence:

1. The first moment of f has components at x such that

$$P^i(x) = \int_{\mathcal{P}_x} p^i F(x, -\lambda(x) p^0) \mu_p = 0, \tag{2.20}$$

because the integrand is antisymmetric in the p^i which vary from $-\infty$ to $+\infty$, while p^0 is positive. The vector P is therefore collinear to the time axis[8] V. The component P^0 in the orthonormal frame,

$$P^0(x) = \int_{\mathcal{P}_x} p^0 F(x, -\lambda(x)p^0)\mu_p, \tag{2.21}$$

is, in the orthonormal frame, the macroscopic particle number density.

2. In the same frame the same antisymmetry reasons imply, $T^{0i} = 0$ and $T^{ij} = 0$ for $i \neq j$. The equality of the components T^{ii}, $i = 1, n$,

$$T^{ii} = \int_{\mathcal{P}_x} (p^i)^2 F(x, -\lambda(x)p^0)\mu_p, \tag{2.22}$$

results from the invariance under p-space rotations of the function F and the volume element μ_p. The component T^{00} is the positive function

$$T^{00} = \int_{\mathcal{P}_x} (p^0)^2 F(x, -\lambda(x)p^0)\mu_p. \tag{2.23}$$

The tensor T is therefore the stress energy tensor of a perfect fluid with flow vector $U = \lambda^{-1}V$, specific energy and pressure given respectively by (2.23) and (2.22). □

Remark 2.5 *The physical interpretation of the first moment that we have given coincides with the interpretation chosen by Marle (1969) and with the Eckart original definition*[9]*. For a discussion and interpretation of the second moment in the general case (dissipative fluids), see the discussion in Marle's paper, pp. 137–43.*

Higher moments are defined as totally symmetric tensors on V given by

$$M^{\alpha_1 \dots \alpha_p} := \int_{\mathcal{P}_x} f(x,p) p^{\alpha_1} \dots p^{\alpha_p} \omega. \tag{2.24}$$

They play an important role in the Müller–Ruggeri extended thermodynamics.

Remark 2.6 *In the case where there are elementary particles with different masses m_a, $a = 1, 2, \dots, N$, there are different phase spaces $\mathcal{P}_{m_a,V}$, and different distribution functions f_a. The moments to consider are the sum of the corresponding moments, for example the second moment is*

$$T^{\alpha\beta}(x) := \sum_a \int_{\mathcal{P}_{m_a},x} f_a(x,p) p^\alpha p^\beta \omega_{m_a,p}. \tag{2.25}$$

[8] Remark that for distribution functions of the assumed form the vector P is timelike, even if the microscopic particles have a zero rest mass.

[9] Eckart, C. (1940) *Phys. Rev.*, **58**, 919–24.

3 Vlasov equations

In the **Vlasov models** one assumes that the gas is so rarefied that the particles trajectories do not cross. Their motion is determined by the average fields they generate on spacetime.

3.1 *Liouville–Vlasov equation.*

We give the name **Liouville–Vlasov** to the equation satisfied by the distribution function in a curved spacetime, in the absence of non-gravitational forces. Then, each particle follows a geodesic of the spacetime metric g. We have seen that the differential system satisfied by a geodesic curve in a pseudo-Riemannian manifold (V, g) reads in local coordinates (the parameter λ is the proper time along the geodesic if it is timelike, a canonical parameter if it is a null geodesic)

$$p^\alpha := \frac{dx^\alpha}{d\lambda}, \quad \frac{dp^\alpha}{d\lambda} = Q^\alpha, \quad \text{with} \quad Q^\alpha := -\Gamma^\alpha_{\lambda\mu} p^\lambda p^m \tag{3.1}$$

where $\Gamma^\alpha_{\lambda\mu}$ are the Christoffel symbols of the metric g. In other words the trajectory of a particle in TV is an element of the geodesic flow generated by the vector field $X = (p, Q)$ whose components p^α, Q^α in a local trivialisation of TV over the domain of a chart of V are given by (3.1).

Exercise. Show by a direct calculation that $X = (p, Q)$ is indeed a vector field on TV

Answer. Let x'^α be another coordinate system in V. Set $A^{\alpha'}_\alpha := \frac{\partial x^{\alpha'}}{\partial x^\alpha}$. The corresponding change of components of a vector in T_xV is $p'^\alpha = \frac{\partial x'^\alpha}{\partial x^\alpha} p^\alpha$. Therefore the change of coordinates $(x^\alpha, p^\alpha) \mapsto (x^{\alpha'}, p^{\alpha'})$ in TV implies

$$\frac{\partial p^{\alpha'}}{\partial p^\alpha} = A^{\alpha'}_\alpha, \quad \frac{\partial p^{\alpha'}}{\partial x^\beta} = \partial_\beta A^{\alpha'}_\alpha. \tag{3.2}$$

Hence for the components $p'^{\alpha'}$, $Q'^{\alpha'}$ of X in the new coordinates on TV

$$p'^{\alpha'} = A^{\alpha'}_\alpha p^\alpha, \quad Q'^{\alpha'} = p^\beta \partial_\beta A^{\alpha'}_\alpha - A^{\alpha'}_\alpha \Gamma^\alpha_{\lambda\mu} p^\lambda p^m.$$

We know that (see Chapter 1)

$$\Gamma'^{\alpha'}_{\beta'\gamma'} = A^{\alpha'}_\alpha \partial_{\beta'} A^\alpha_{\gamma'} + A^{\alpha'}_\alpha A^\beta_{\beta'} A^\gamma_{\gamma'} \Gamma^\alpha_{\beta\gamma}. \tag{3.3}$$

Some straightforward manipulations shows that

$$Q'^{\alpha'} \equiv -\Gamma'^{\alpha'}_{\beta'\gamma'} p^{\beta'} p^{\gamma'}. \tag{3.4}$$

The following lemma is another formulation of a property already seen in other contexts.

Lemma 3.1 *The scalar $g(p, p)$ is constant along an orbit of X in TV.*

Proof The equality

$$\left(p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu \frac{\partial}{\partial p^\alpha}\right) g_{\sigma\rho} p^\sigma p^\rho = 0, \tag{3.5}$$

when x and p are independent variables, is a consequence of the values of the Christoffel symbols. □

This property of $g(p,p)$ implies that if a particle has rest mass m its trajectory lies in $\mathcal{P}_m$.

Theorem 3.2 *(Liouville theorem) The volume form θ is invariant under the geodesic flow; that is, $\mathcal{L}_X$ denoting the Lie derivative with respect to X, it holds that*

$$\mathcal{L}_X \theta = 0. \tag{3.6}$$

Proof In local coordinates $(x^A) := (x^\alpha, p^{\bar\alpha})$ on TV, the Lie derivative[10] of the exterior $2(n+1)$-form θ with respect to the vector field $X = (p, Q)$, *i.e.* $(X^A) = (p^\alpha, Q^{\bar\alpha})$, reads, using the fact that the components of θ do not depend on p,

$$(\mathcal{L}_X\theta)_{01\ldots nn+1\ldots 2(n+1)} = p^\alpha \frac{\partial}{\partial x^\alpha}\theta_{01\ldots nn+1\ldots 2(n+1)} + \tag{3.7}$$

$$\theta_{A1\ldots 2(n+1)} \frac{\partial X^A}{\partial x^0} + \theta_{0A\ldots 2(n+1)} \frac{\partial X^A}{\partial x^1} + \cdots + \theta_{01\ldots 2n+1A} \frac{\partial X^A}{\partial x^{2(n+1)}} \tag{3.8}$$

The expression of θ gives that

$$\frac{\partial}{\partial x^\alpha}\theta_{01\ldots nn+1\ldots 2(n+1)} = g^{\lambda\mu}\left(\frac{\partial}{\partial x^\alpha} g_{\lambda\mu}\right)\theta_{01\ldots nn+1\ldots 2(n+1)}; \tag{3.9}$$

on the other hand, since θ is antisymmetric the same index cannot appear twice in its components, therefore the second line is equal to (recall that the components p^α of X do not depend on x^α)

$$\theta_{01\ldots 2(n+1)} \frac{\partial X^A}{\partial x^A} = \theta_{01\ldots 2(n+1)} \frac{\partial Q^\alpha}{\partial p^\alpha}. \tag{3.10}$$

The expression of Q gives that

$$\frac{\partial Q^\alpha}{\partial p^\alpha} = -2\Gamma^\lambda_{\lambda\alpha} p^\alpha. \tag{3.11}$$

The result $\mathcal{L}_X\theta = 0$ follows from the expression of the Christoffel symbols □

[10] CB-DM1 III C.

Corollary 3.3

1. *The* $2n+1$*-exterior form*[11] *(CB-DM1 IV A)* $i_X\theta$ *is a closed form, i.e.* $di_X\theta = 0$.
2. *The form* $i_X\theta$ *is invariant under the flow of any vector field of the type* aX*, where* a *is a positive scalar function on* TV.

Proof

1. The volume form θ is closed, since it is of maximum degree on TV. The property $di_X\theta = 0$ is a consequence of the previous theorem and of the identity valid for the Lie derivative of any exterior form

$$\mathcal{L}_X \equiv di_X + i_X d. \tag{3.12}$$

2. Since $i_X\theta$ is closed its Lie derivative with respect to aX reduces to

$$\mathcal{L}_{aX}\ i_X\theta \equiv d(i_{aX}\ i_X\theta). \tag{3.13}$$

but acting on an exterior form we have $i_{aX}\ i_X \equiv a(i_X)^2 \equiv 0$. □

The corollary shows that $i_X\theta$ is an absolute integral invariant[12] for the trajectories of X, that is the integrals of $i_X\theta$ on a $2n+1$-dimensional domain D of TV transversal to the trajectories of X and on the image D' of D under the local 1-parameter group generated by a vector field aX are equal. It is legitimate to consider $i_X\theta$ as a measure element for the number of trajectories[13] of X.

By the definition $(dx^A) := (dx^\alpha, dp^{\bar\alpha}) = Xd\lambda$ of the vector field X, the integral of a 1-form τ on a trajectory γ of X is

$$\int_\gamma i_X\tau d\lambda, \quad i_X\tau \equiv X^A\tau_A \equiv p^\alpha\tau_\alpha + Q^{\bar\alpha}\tau_{\bar\alpha}. \tag{3.14}$$

This integral defines the proper time (or canonical parameter) on γ, up to choice of origin, if and only if it is such that $i_X\tau \equiv 1$. For such a 1-form τ we have

$$\tau \wedge i_X\theta = \tau_A dx^A \wedge (-1)^A X^A \theta_{0\ldots A\ldots 2(n+1)} dx^0 \wedge \ldots \wedge d\hat{x}^A \wedge \ldots dx^{2(n+1)} = \theta. \tag{3.15}$$

This equality makes coherent the interpretations of the products $f\theta$ and $fi_X\theta$, with f the distribution function, respectively as the presence number density and the density of trajectories.

Analogous definitions and properties hold for particles with a given rest mass.

[11] d is the exterior derivative and i the interior product given for a p form by the $p-1$ form of totally antisymmetric components $(i_X\theta)_{A_2\ldots A_p} = X^A\theta_{AA_2\ldots A_p}$.

[12] See CB-DM1 IV C.

[13] The trajectories of X and aX are the same geometrical curves, but the one-parameter transformation group they generate depend on the parametrization of these curves, i.e. on the scalar function a.

In a collisionless model the physical law of conservation of particles forces the form $f\theta$ to be invariant under the vector field X and hence to have a zero Lie derivative with respect to X. Since we already know that $\mathcal{L}_X\theta = 0$ the invariance reduces to the **Liouville–Vlasov** equation for the distribution function f

$$\mathcal{L}_X f \equiv p^\alpha \frac{\partial f}{\partial x^\alpha} + Q^\alpha \frac{\partial f}{\partial p^\alpha} = 0. \tag{3.16}$$

which says that in phase space the derivative of the distribution function in the direction of X vanishes.

3.2 *Maxwell–Vlasov equation*

When, in addition to gravitation, the particles are subjected to some other force, represented by a vector Φ tangent to the spacetime V, the vector X tangent to the particles trajectories in the phase space $\mathcal{P}_V$ over a spacetime (V, g) is in a trivialization of $\mathcal{P}_V$

$$(X^A) = (p^\alpha, Q^\beta + \Phi^\beta). \tag{3.17}$$

Theorem 3.4 *Lemma (3.1) and Theorem (3.2) hold when Φ is the Lorentz force of an electromagnetic field.*

Proof The vector X is the sum of the gravitational vector (p, Q) and the vector $(0, \Phi)$. The Lie derivative with respect to X is the sum of the derivatives with respect to these two vectors. We already know that the first one is zero for the function $g(p, p)$ and the volume form θ. The Lorentz force of an electromagnetic field F on particles with electric charge e is

$$\Phi^\alpha := eF^{\alpha\beta}p_\beta \tag{3.18}$$

The Lie derivative of $g(p, p)$ with respect to $(0, \Phi)$ reduces to

$$\Phi^\alpha \frac{\partial g(p,p)}{\partial p^\alpha} \equiv 2eF^{\alpha\beta}p_\beta p_\alpha = 0 \tag{3.19}$$

due to the antisymmetry of F.

We have

$$i_{(0,\Phi)}\theta = \omega_x \wedge i_\Phi \omega_p \tag{3.20}$$

Straightforward calculation shows that

$$d i_\Phi \omega_p \equiv \frac{\partial \Phi^\alpha}{\partial p^\alpha}\omega_p = 0, \tag{3.21}$$

because $\frac{\partial \Phi^\alpha}{\partial p^\alpha} = 0$ if Φ is given by (3.18). □

Remark. *The proof shows that the theorem holds, more generally, if Φ is orthogonal to p and divergence free with respect to p.*

3.3 *Yang–Mills–Vlasov equation*

The Yang–Mills–Vlasov equation is analogous to the Maxwell–Vlasov equation, but the electromagnetic field is replaced by a Yang–Mills field taking its values in a Lie algebra $\mathcal{G}$ and the electric charge e, a constant, is replaced by a function q on the spacetime V with values in the Lie algebra $\mathcal{G}$. The phase space for the kinetic theory is now the product $\mathcal{P}_V \times \mathcal{G}$. The trajectory of a particle in this phase space is a solution of the differential system

$$\frac{dx^\alpha}{ds} = p^\alpha, \quad \frac{dp^\alpha}{ds} = Q^\alpha + q.F^{\alpha\beta}p_\beta, \quad \frac{dq^a}{ds} = \widehat{Q}^\alpha := -p^\alpha[A_\alpha, , q]^a,$$

$[.,.]$ is the bracket in the Lie algebra $\mathcal{G}$, i.e. $[A_\alpha, , q]^a := c^a_{bc}A^b_\alpha q^c$. The distribution function f is now function of x, p, q. In the absence of collisions and other forces f satisfies the Yang–Mills–Vlasov equation

$$\mathcal{L}_X f \equiv p^\alpha \frac{\partial f}{\partial x^\alpha} + (Q^\alpha + q.F^{\alpha\beta}p_\beta)\frac{\partial f}{\partial p^\alpha} + \widehat{Q}^a \frac{\partial f}{\partial q^\alpha} = 0. \tag{3.22}$$

3.4 *Particles of a given rest mass*

In application to galaxies, or the whole cosmos, it is reasonable to assume that the masses of the "particles" vary between two positive numbers. The case of particles having a given rest mass appears in relativistic plasmas. We consider in this section, as an example, a model with N particles of different kinds, with rest masses m_I and electric charges e_I. We denote by $f_I(x,p)$ their distribution functions on TV. Each f_I satisfies a Maxwell–Vlasov equation

$$\mathcal{L}_{X_I} f_I = 0, \quad X_I := (p, Q + \Phi_I), \quad \Phi^\alpha_I := e_I F^{\alpha\beta}p_\alpha \tag{3.23}$$

on the phase space $\mathcal{P}_{V,m_I}$ of a particle of kind I. The corresponding momenta lie in the mass hyperboloid

$$g_{\alpha\beta}p^\alpha p^\beta = -m_I^2. \tag{3.24}$$

We take the p^i as coordinates on this mass hyperboloid. The distribution function f_{m_I} in such coordinates is

$$f_{m_I}(x^\alpha, p^i) = f_I(x^\alpha, p^i, p^0(p^i)) \tag{3.25}$$

We deduce from (3.24) that

$$\frac{\partial p^0}{\partial p^i} = -\frac{p_i}{p_0}, \quad \frac{\partial p^0}{\partial x^\alpha} = -\frac{p^\lambda p^\mu}{2p_0}\frac{\partial g_{\lambda\mu}}{\partial x^\alpha}. \tag{3.26}$$

hence, with $\Psi := Q + \Phi$,

$$p^\alpha \frac{\partial f_{m_I}}{\partial x^\alpha} + \Psi^i \frac{\partial f_{m_I}}{\partial p^i} = p^\alpha \left(\frac{\partial f_I}{\partial x^\alpha} - \frac{p^\lambda p^\mu}{2p_0}\frac{\partial g_{\lambda\mu}}{\partial x^\alpha}\frac{\partial f_I}{\partial p^0} \right) + \Psi^i \left(\frac{\partial f_I}{dp^i} - \frac{p_i}{p_0}\frac{\partial f_I}{\partial p^0} \right). \tag{3.27}$$

This identity together with identity

$$Q^\alpha p_\alpha + \frac{p^\alpha p^\lambda p^\mu}{2p_0}\frac{\partial g_{\lambda\mu}}{\partial x^\alpha} \equiv 0, \tag{3.28}$$

and the equation satisfied by the force Φ

$$\Phi^\alpha p_\alpha = 0 \tag{3.29}$$

show that

$$p^0\frac{\partial f_{m_I}}{\partial x^0} + p^i\frac{\partial f_{m_I}}{\partial x^i} + \Psi^i\frac{\partial f_{m_I}}{\partial p^i} \equiv p^\alpha\frac{\partial f_I}{\partial x^\alpha} + Q^\alpha\frac{\partial f_I}{\partial p^i} + \Phi^\alpha\frac{\partial f_I}{dp^\alpha} \equiv \mathcal{L}_X f_I.. \tag{3.30}$$

The Maxwell–Vlasov equation (3.23) for particles of a given rest mass reduces therefore to the equation

$$\frac{\partial f_{m_I}}{\partial x^0} + \frac{p^i}{p^0}\frac{\partial f_{m_I}}{\partial x^i} + \frac{\Psi^i}{p^0}\frac{\partial f_{m_I}}{\partial p^i} = 0, \quad \text{with} \quad \Psi^i \equiv -\Gamma^i_{\lambda\mu}p^\lambda p^\mu + \Phi^i_I. \tag{3.31}$$

3.5 *Conservation of moments*

The following theorem, important in the case of the second moment for consistency with the Einstein equations, holds for all the moments in the purely gravitational case. For the first moment it expresses the macroscopic conservation law of matter.

Theorem 3.5 *If the distribution f satisfies the Liouville–Vlasov equation (3.16) then its moments satisfy the conservation laws*

$$\nabla_{\alpha_1} M^{\alpha_1\alpha_2\ldots\alpha_p} = 0 \tag{3.32}$$

where ∇ is the covariant derivative in the spacetime metric g.

Proof Choosing at the point x coordinates such that the first derivatives $\partial_\alpha g_{\lambda\mu}$, hence the Christoffel symbols, vanish at that point one finds immediately the following relation between tensors on V:

$$\nabla_{\alpha_1} M^{\alpha_1\alpha_2\ldots\alpha_p} = \int_{\mathcal{P}_x} (\mathcal{L}_X f)(x,p)p^{\alpha_2}\ldots p^{\alpha_p}\mu_p, \tag{3.33}$$

since in such coordinates covariant derivatives reduce at the point x to ordinary partial derivatives, while $\mathcal{L}_X f$ reduces to $p^\alpha(\partial f/\partial x^\alpha)$ and $\partial_\alpha(\det g)^{\frac{1}{2}}$ vanishes. □

Exercise. Prove the equation in arbitrary coordinates in the case of particles of a given rest mass.

Hint. Write the covariant divergence of the tensor M on V, computing its partial derivative using the fact that $\mu_p \equiv (\det g)^{\frac{1}{2}}d^{n+1}p$, the expression of the Lie derivative and an integration by parts.

4 Cauchy problem for the Liouville–Vlasov equation

Let the $n+1$-dimensional spacetime (V, g) be a given Lorentzian manifold. The Liouville equation is a first-order linear partial differential equation on the tangent bundle TV. Let $\bar{M}$ be an n-dimensional submanifold of V and $\mathcal{P}_{\bar{M}}$, the subbundle of TV with basis $\bar{M}$. The Cauchy data is a function $\bar{f}$ on $\mathcal{P}_{\bar{M}}$, i.e. a function $\bar{f}(\xi, \pi)$ with $\xi \in \bar{M}$, $\pi \in \mathcal{P}_\xi$ the fibre at ξ.

4.1 *General solution*

The Liouville–Vlasov equation can be written as an ordinary differential equation on the phase space $\mathcal{P}_V$ corresponding to the considered family of particles,

$$\frac{df(x(\lambda), p(\lambda))}{d\lambda} = 0, \tag{4.1}$$

by solving the characteristic system

$$\frac{dx^\alpha}{d\lambda} = p^\alpha, \quad \frac{dp^\alpha}{d\lambda} = Q^\alpha \equiv -\Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu. \tag{4.2}$$

This quasilinear first-order differential system has, for λ small enough, one and only one solution taking for $\lambda = 0$ given values ξ^α, π^α if its coefficients are Lipschitzian in x and p, hence if the spacetime metric is $C^{1,1}$ (i.e. C^1 with Lipschitzian derivatives); the mapping $(\xi, \pi) \mapsto (x, p)$ defined for each λ by this solution is C^1 if the coefficients of the differential system are C^1, i.e. if the spacetime metric is C^2.

Assume given in V an initial submanifold M_0 with equation $x^0 = 0$. Denote by

$$x^\alpha(\lambda, \xi^i, \pi^\alpha), \quad p^\alpha(\lambda, \xi^i, \pi^\alpha) \tag{4.3}$$

the solution of the differential system (4.2) taking on M_0, for $\lambda = 0$, the values

$$\xi^0 = x^0(0, \xi^i, \pi^\alpha) = 0, \quad \xi^i = x^i(0, \xi^i, \pi^\alpha), \quad \pi^\alpha = p^\alpha(0, \xi^i, \pi^\alpha). \tag{4.4}$$

The differentiable mapping $(\lambda, \xi^i, \pi^\alpha) \mapsto (x^\alpha(\lambda, \xi^i, \pi^\alpha), p^\alpha(\lambda, \xi^i, \pi^\alpha))$ reduces to the identity for $\lambda = 0$; it is therefore invertible for small enough λ. The inverse mapping gives, using Equation (4.1), the following solution of the Liouville–Vlasov equation in a neighbourhood of M_0

$$f(x, p) \equiv \bar{f}(\xi(x, p), \pi(x, p)). \tag{4.5}$$

The support of f in $\mathcal{P}_V$ is covered by trajectories of X issued from $\mathcal{P}_{M_0}$. Its boundary in TV is generated by trajectories of X issued from the boundary of $\mathcal{P}_{M_0}$ in TV.

4.2 *Distribution function on a Robertson–Walker space time*

Consider for example a spacetime $M \times R$ with a Robertson–Walker metric[14]

$$g \equiv -dt^2 + R^2(t)\sigma^2 \quad \text{with} \quad \sigma^2 \equiv \gamma_{ij} dx^i dx^j$$

[14] See computation of Christoffel symbols in Chapter 5.

σ is a given Riemannian metric on M. Straightforward computation using $Q^\alpha \equiv -\Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu$ gives that

$$\mathcal{L}_X f \equiv p^\alpha \frac{\partial f}{\partial x^\alpha} - RR'\gamma_{ij}p^i p^j \frac{\partial f}{\partial p^0} - 2R^{-1}R'p^0 p^i \frac{\partial f}{\partial p^i}.$$

We look for a solution depending only on t and p^0, for particles of a given rest mass m, that is when

$$R^2\gamma_{ij}p^i p^j = (p^0)^2 - m^2,$$

the equation for f_m reduces to

$$p^0 \frac{\partial f_m}{\partial t} - R^{-1}R'\{(p^0)^2 - m^2\}\frac{\partial f_m}{\partial p^0} = 0.$$

Taking R instead of t as a variable the equation reads as the following linear first-order partial differential equation

$$R\frac{\partial f_m}{\partial R} - \{(p^0)^2 - m^2\}\frac{1}{p^0}\frac{\partial f_m}{\partial p^0} = 0.$$

The general solution is constant along the rays (bicharacteristics) which satisfy the differential system

$$\frac{dR}{R} = -\frac{p^0 dp^0}{(p^0)^2 - 1} = d\lambda.$$

These rays are such that

$$logR + \frac{1}{2}log((p^0)^2 - m^2) = \text{constant}, \quad \text{i.e. } R^2\{(p^0)^2 - m^2\} = \text{constant}.$$

The distribution f_m is therefore an arbitrary function of the scalar $R^2\{(p^0)^2 - m^2\}$. Suppose for instance that f_m vanishes at time t_0 for particles with momentum such that

$$(p^0)^2 \geq m^2 + KR^{-2}(t_0),$$

then at time t the function f_m vanishes at time t for particles with momentum

$$(p^0)^2 \geq m^2 + KR^{-2}(t).$$

Hence the maximum of possible energy p^0 of particles with a given rest mass decreases with expansion, as foreseen physically.

4.3 *Energy estimates*

Our aim is to study the Einstein–Vlasov system. Therefore the integrals (2.17) defining the stress energy tensor T of the distribution f must have a meaning as a source for the Einstein equations. In the next section we prove a local existence theorem for a solution of the Cauchy problem for the coupled Einstein–Vlasov system, using general methods applying to hyperbolic equations, that is

energy estimates. We have seen in Chapter 6 that the Cauchy problem for the Einstein equations in wave gauge has a local in time solution for Cauchy data in $H_s^{loc} \times H_{s-1}^{loc}$, $s > \frac{n}{2} + 1$ if the stress energy tensor is in H_{s-1}^{loc}. In this subsection we establish estimates for the stress energy tensor of a distribution function f satisfying a Liouville–Vlasov equation through energy-type estimates of f on a phase space $\mathcal{P}_V$ subbundle of the tangent space TV of a given Lorentzian manifold (V, g).

We assume, as in Chapter 6, that (V, g) is a sliced Lorentzian manifold that is $V = M \times R$ and

$$g \equiv -N^2 dt^2 + g_{ij}\theta^i\theta^j, \quad \theta^i \equiv dx^i + \beta^i dt. \tag{4.6}$$

To define Sobolev spaces independent of the metric g, which will also be an unknown in the next section, we introduce on M as in previous chapters a smooth Sobolev regular Riemannian metric e and on V the Riemannian metric $e + dt^2$. Since $V = M \times R$, $TV = TM \times TR$,

$$\hat{e} = e_{ij}dx^i dx^j + dt^2 + e_{ij}dp^i dp^j + (dp^0)^2 \tag{4.7}$$

is a Sobolev regular Riemannian metric on TV. The intrinsic formulations on a manifold are somewhat complicated, though they do not involve conceptual difficulties. For simplicity we treat the case where the space manifold M is diffeomorphic to R^n and we choose for e the Euclidean metric. We suppose that on a spacetime slice $V_T := M \times [0, T]$ the induced metric g_t on M_t is uniformly equivalent to e, the lapse N is uniformly equivalent to 1 and the shift β uniformly bounded, all these quantities being continuous[15]. We suppose that $\partial g \in \cap_{k=0,1} C^k([0, T], H_{s-k})$, where H_s is the usual space of tensors on (M, e) defined in Appendix I, $s > \frac{n}{2} + 1$, so g is C^2. We leave to the reader the proof of variants of the obtained results under variants of our hypotheses.

Standard results for hyperbolic equations do not apply directly to the Vlasov equation on $\mathcal{P}_V$, because its coefficients are unbounded. We treat explicitly[16] the case most relevant for General Relativity of particles with masses bounded below and above by positive numbers, m and M. The $\mathcal{P}_{x,t}$ fibre at $(x, t) \in M \times R$ of the phase space $\mathcal{P}_V$ is then

$$\mathcal{P}_{x,t} \equiv T_{x,t}V \cap \{0 < m^2 \leq N^2(p^0)^2 - g_{ij}p^i p^j \leq M^2,\ p^0 > 0\}. \tag{4.8}$$

We denote by $\mathcal{P}_{V_T} := \hat{V}_T$ the subbundle of $\mathcal{P}_V$ with basis V_T. It is a manifold with boundary $\partial\hat{V}_T$, a boundary made of manifolds with opposite orientations $\hat{M}_T := \mathcal{P}_{M_T}$ and $\hat{M}_0 := \mathcal{P}_{M_0}$ and of a lateral boundary L_T, the bundle on V_T with fibres $p^0 > 0$ and either

$$m^2 = N^2(p^0)^2 - g_{ij}p^i p^j \quad \text{or} \quad N^2(p^0)^2 - g_{ij}p^i p^j = M^2. \tag{4.9}$$

[15] Such manifolds were called regularly sliced in Chapter 6.

[16] As occurs in cosmology. For particles of given mass see for instance Choquet-Bruhat, Y. and Bancel, D. (1973) *Commun. Math. Phys.*, **33**, 1–14.

4.3.1 *Fundamental estimate*

Let h be a C^1 function on $\mathcal{P}_V$. Since $hf^2\theta$ is, like θ, an $n+1$-form on an $n+1$ manifold, its exterior differential is identically zero and the expression (3.12) of a Lie derivative implies

$$\mathcal{L}_X(hf^2\theta) \equiv di_X(hf^2\theta).$$

Therefore, if the Stokes formula applies, it holds that:

$$\int_{\hat{V}_T} \mathcal{L}_X(hf^2\theta) = \int_{\partial\hat{V}_T} i_X(hf^2\theta) \equiv \int_{\partial\hat{V}_T} hf^2 i_X\theta. \tag{4.10}$$

Since the volume form θ is invariant under the flow of X it holds that, if f satisfies the Liouville–Vlasov equation,

$$\mathcal{L}_X(hf^2\theta) = (\mathcal{L}_X h)f^2\theta. \tag{4.11}$$

On the submanifold $\hat{M}_t$ we have $dt \equiv dx^0 = 0$. Therefore the expression of the vector X gives

$$i_X\theta = p^0 N_t^2 \omega_{\bar{g}_t} \wedge \omega_{p,\bar{g}_t}, \quad \text{on} \quad \hat{M}_t. \tag{4.12}$$

where $\omega_{\bar{g}_t}$ is the volume form of $(M, \bar{g}_t)$ and $\omega_{p,\bar{g}_t} = (det\bar{g}_t)^{1/2} dp^0 \wedge \cdots \wedge dp^n$.

The lateral boundary L_T is generated by trajectories of the vector field X; hence by the definition of i_x,

$$i_X\theta = 0 \quad \text{on} \quad L_T. \tag{4.13}$$

Exercise. *Prove this formula by computations analogous to the computation made in Section 3.4 for particles with a given rest mass, using the identity valid on* L_T

$$2g_{\alpha\beta}p^\beta dp^\alpha + dg_{\alpha\beta}p^{\alpha\beta} = 0. \tag{4.14}$$

We are led to the following definition.

Definition 4.1 *The* ***h-energy*** *of* f *on* $\hat{M}_t$ *is the value of the integral of the exterior form* $hf^2 i_X\theta$ *on the submanifold* $\hat{M}_t$ *with orientation induced by the positive time orientation of spacetime*[17]*, that is we set*

$$E_h(t) := \int_{\hat{M}_t} hf^2 i_X\theta \equiv \int_{\hat{M}_t} hf^2 p^0 N_t \mu_{g_t}\mu_p, \tag{4.15}$$

$$\mu_{g_t} = (\det \bar{g}_t)^{\frac{1}{2}} dx^1 \dots dx^n, \quad \mu_p = (\det g)^{\frac{1}{2}} dp^0 dp^1 \dots dp^n. \tag{4.16}$$

The formulas (4.10) and (4.11) give

$$E_h(T) - E_h(0) = \int_0^T \int_{\hat{M}_t} (\mathcal{L}_X h) f^2 N_t \mu_{\bar{g}_t}\mu_p dt. \tag{4.17}$$

[17] For integration of exterior forms on oriented manifolds see for instance see CB-DM1 IV.

To obtain L^2 bounds of the stress energy tensor of the distribution f it is convenient to introduce a weight of the form $h = (v_\alpha p^\alpha)^P$, P some number.

Lemma 4.2 *Let $h := (v_\alpha p^\alpha)^P$ with v a past-oriented timelike vector. Then:*

1. *$\mathcal{L}_X h = 0$ if and only if v is a Killing vector of the spacetime metric g.*
2. *In the general case we choose for v the tangent to the time line*[18]*, i.e. $v^\alpha = -\delta^\alpha_0$, $h = (p^0)^P$. Under the hypothesis made on the metric there exists a constant C such that*

$$|\mathcal{L}_X h| \leq C h p^0. \tag{4.18}$$

Proof

1. Since h is a scalar, its Lie derivative is its ordinary derivative in the direction of X

$$\mathcal{L}_X h = P(v_\alpha p^\alpha)^{P-1} X^A \partial_A (v_\lambda p^\lambda) \tag{4.19}$$

with, using the expression of X and changing names of indices,

$$X^A \partial_A (v_\lambda p^\lambda) = p^\mu p^\lambda \left(\frac{\partial}{\partial x^\mu} v_\lambda - v_\alpha \Gamma^\alpha_{\lambda\mu} \right) \equiv p^\mu p^\lambda \nabla_\mu v_\lambda. \tag{4.20}$$

Therefore, using the expression

$$(\mathcal{L}_v g)_{\mu\lambda} \equiv \nabla_\mu v_\lambda + \nabla_\lambda v_\mu \tag{4.21}$$

of the Lie derivative of g with respect to v,

$$\mathcal{L}_X h = \frac{1}{2} P (v_\alpha p^\alpha)^{P-1} p^\mu p^\lambda (\mathcal{L}_v g)_{\mu\lambda}. \tag{4.22}$$

Hence $\mathcal{L}_X h = 0$ if v is a Killing vector of g.

2. When $h = (p^0)^P$ we have

$$\mathcal{L}_X h = -P (p^0)^{P-1} p^\lambda p^\mu \Gamma^0_{\lambda\mu}, \tag{4.23}$$

the result (4.18), with

$$C = \underset{V_T}{Sup} \left| P \frac{p^\lambda p^\mu}{(p^0)^2} \Gamma^0_{\lambda\mu} \right|, \tag{4.24}$$

follows from the hypothesis made on g and the inequality (4.8) satisfied by p.

□

[18] The choice of another strictly timelike vector will give equivalent results, with more cumbersome writing.

Theorem 4.3 *Let $h := (v_\alpha p^\alpha)^P$ with v a past-oriented timelike vector, then if f satisfies the Liouville–Vlasov equation:*

1. *The h weighted energy of f is conserved in time if v is a Killing vector of the spacetime.*

2. *If v is the tangent to the time line*[19]*, i.e. $v^\alpha = \delta^\alpha_0$, $h = (p^0)^n$. Under the hypothesis made on the metric there exists a constant C such that the h-energy of f satisfies the* ***energy inequality***

$$E_h(t) \leq E_h(0)e^{Ct}. \tag{4.25}$$

Proof

1. If $\mathcal{L}_v g = 0$, then $\mathcal{L}_X h = 0$ and Equation (4.17) implies

$$E_h(t) = E_h(0). \tag{4.26}$$

2. The energy equality (4.17) and the bound (4.18) imply the inequality

$$E_h(t) \leq E_h(0) + C\int_0^t E_h(\tau)d\tau, \tag{4.27}$$

hence, by the Gromwall lemma, the **energy inequality** (4.25). □

4.3.2 *Estimate of ∂f*

Before generalizing Definition 4.1 and the estimate (4.27) to derivatives of f, as needed for estimates in the coupling with Einstein's equations, we study the Lie derivatives of the partial derivatives of f. We find that Lie derivatives of partial derivatives of f with respect to x and with respect to p have a different asymptotic behaviour on the fibres of the phase space.

We denote by $|\partial_x f|$ the scalar on $\mathcal{P}_V$, norm in the Euclidean metric $(e^{\alpha\beta})$ of the gradient of f with respect to the spacetime coordinates x^α

$$|\partial_x f|^2 := e^{\alpha\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}; \tag{4.28}$$

we denote by $|\partial_p f|$, the Euclidean norm of the gradient of f in the fibre

$$|\partial_p f|^2 := e^{\alpha\beta}\frac{\partial f}{\partial p^\alpha}\frac{\partial f}{\partial p^\beta}. \tag{4.29}$$

Lemma 4.4 *If the function f satisfies the Liouville–Vlasov equation $\mathcal{L}_X f = 0$, then the scalars $|\partial_x f|^2$ and $|\partial_p f|^2$ are such that*

$$\mathcal{L}_X|\partial_x f|^2 = q_x(\partial f, \partial f), \quad \mathcal{L}_X|\partial_p f|^2 = q_p(\partial f, \partial f) \tag{4.30}$$

[19] The choice of another strictly timelike vector will give equivalent results, with more cumbersome writing.

where q_x and q_p are quadratic forms in the gradient of f on the phase space $\mathcal{P}_V$. If the spacetime metric is uniformly C^2 on the strip V_T then q_x and q_p satisfy on $\mathcal{P}_{V_T}$ inequalities of the form

$$|q_x(\partial f, \partial f)| \leq C\{(p^0)^2|\partial_x f|\ |\partial_p f|, \tag{4.31}$$

$$|q_p(\partial f, \partial f)| \leq C(|\partial_x f|\ |\partial_p f| + p^0|\partial_p f|^2 \tag{4.32}$$

Proof Since the $e^{\alpha\beta}$ are constants we have

$$\mathcal{L}_X|\partial_x f|^2 \equiv 2e^{\alpha\beta}\frac{\partial f}{\partial x^\alpha}X^A\frac{\partial}{\partial x^A}\frac{\partial}{\partial x^\beta}f. \tag{4.33}$$

Using commutation of partial derivatives we find

$$\mathcal{L}_X|\bar{\partial} f|^2 \equiv 2e^{\alpha\beta}\frac{\partial f}{\partial x^\alpha}\left\{\frac{\partial}{\partial x^\beta}\left(X^A\frac{\partial f}{\partial x^A}\right) - \frac{\partial X^A}{\partial x^\beta}\frac{\partial f}{\partial x^A}\right\}. \tag{4.34}$$

Therefore when $\mathcal{L}_X f = 0$ we find that $\mathcal{L}_X|\partial_x f|^2$ is a quadratic form in ∂f, namely:

$$\mathcal{L}_X|\partial_x f|^2 = q_x(\partial f, \partial f) \equiv -2e^{\alpha\beta}\frac{\partial Q^\lambda}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}\frac{\partial f}{\partial p^\lambda}. \tag{4.35}$$

We consider the scalar $|\partial_p f|$, the Euclidean norm of the gradient of f in the fibre. An analogous computation gives

$$\mathcal{L}_X|\partial_p f|^2 \equiv 2e^{\alpha\beta}\frac{\partial f}{\partial p^\alpha}\left\{\frac{\partial}{\partial p^\beta}\left(X^A\frac{\partial f}{\partial x^A}\right) - \frac{\partial X^A}{\partial p^\beta}\frac{\partial f}{\partial x^A}\right\} \tag{4.36}$$

hence, if f satisfies the Liouville–Vlasov equation

$$\mathcal{L}_X|\partial_p f|^2 = q_p(\partial f, \partial f) := -2e^{\alpha\beta}\frac{\partial f}{\partial p^\alpha}\left\{\frac{\partial f}{\partial x^\beta} + \frac{\partial Q^\gamma}{\partial p^\beta}\frac{\partial f}{\partial p^\gamma}\right\}. \tag{4.37}$$

In the purely gravitational case it holds that:

$$Q^\alpha = -\Gamma^\alpha_{\lambda\mu}p^\lambda p^\mu, \quad \text{hence} \quad \frac{\partial Q^\alpha}{\partial x^j} = -p^\lambda p^\mu\frac{\partial\Gamma^\alpha_{\lambda\mu}}{\partial x^j}, \quad \frac{\partial Q^\alpha}{\partial p^\gamma} = -2p^\mu\Gamma^\alpha_{\gamma\mu}. \tag{4.38}$$

Therefore if the metric g is uniformly C^2 there exists a number C depending only on its C^2 bound such that the quadratic forms q_x and q_p satisfy the given inequalities. □

The lemma leads to a choice of different weights for $\partial_p f$ and $\partial_x f$ in the following definition.

Definition 4.5 *The P-energy of ∂f at time t is the integral on $\hat{M}_t$*

$$E^{(1)}_P(t) := \int_{\hat{M}_t}\{(p^0)^P|\partial_x f|^2 + (p^0)^{P+2}|\partial_p f|^2\}N\mu_{\bar{g}}\mu_p, \tag{4.39}$$

Theorem 4.6 *If the spacetime metric g is uniformly C^2 on the strip V_T the P-energy satisfies an inequality*

$$E_P^{(1)}(t) \leq E_P^{(1)}(0)e^{Ct}, \quad t \in [0,T], \tag{4.40}$$

where C is a constant depending only on the C^2 bounds of g on V_T

Proof The function $E_P^{(1)}(t)$ is the integral on the positively oriented manifold $\hat{M}_t$ of the sum of two $2n+1$ exterior forms

$$(p^0)^{P-1}|\partial_x f|^2 i_X\theta + (p^0)^{P'-1}|\partial_p f|^2 i_X\theta, \quad P' = P+2. \tag{4.41}$$

We have seen that this geometric interpretation implies that

$$E_P^{(1)}(t) - E_P^{(1)}(0) = \int_0^t \int_{\hat{M}_\tau} \mathcal{L}_X\{(p^0)^{P-1}|\partial_x f|^2 + (p^0)^{P'-1}|\partial_p f|^2\}\theta \tag{4.42}$$

For any pair of scalars F and h on the bundle $\mathcal{P}_V$ it holds that

$$\mathcal{L}_X(hF\theta) = \{(\mathcal{L}_X h)F + h(\mathcal{L}_X F)\}\theta. \tag{4.43}$$

Therefore

$$\mathcal{L}_X\{(p^0)^{P-1}|\partial_x f|^2\} = |\partial_x f|^2\mathcal{L}_X(p^0)^{P-1} + (p^0)^{P-1}\mathcal{L}_X|\partial_x f|^2 \tag{4.44}$$

with, using previous results,

$$\mathcal{L}_X(p^0)^{P-1} = (P-1)(p^0)^{P-2}\mathcal{L}_X p^0 \leq C(p^0)^P \tag{4.45}$$

and

$$(p^0)^{P-1}|\mathcal{L}_X|\partial_x f|^2| \leq C(p^0)^{P/2}|\partial_x f|\ (p^0)^{P/2+1}|\partial_p f| \tag{4.46}$$

On the other hand, with similar arguments,

$$\mathcal{L}_X\{(p^0)^{P'-1}|\partial_p f|^2|\} \leq C\{(p^0)^{P'}|\partial_p f|^2 + (p^0)^{P'-1}|\partial_p f|\ \{|p^0|\partial_p f| + |\partial_x f|\} \tag{4.47}$$

we set $P' = P+2$ and write the above inequality

$$\mathcal{L}_X\{(p^0)^{P'-1}|\partial_p f|^2|\} \leq C\{(p^0)^{P'}|\partial_p f|^2 + (p^0)^{P'/2}|\partial_p f|\ (p^0)^{P/2}|\partial_x f|\}. \tag{4.48}$$

The equality (4.42) implies then the inequality

$$E_P^{(1)}(t) - E_P^{(1)}(0) \leq C\int_0^t E_P^{(1)}(\tau)d\tau. \tag{4.49}$$

The inequality (4.40) follows from the Gromwall lemma. □

4.3.3 *Estimates of higher derivatives of f*

Estimates of weighted L^2 norms of higher derivatives of f are obtained by a similar method. They are obtained through bounds of the Lie derivatives of pointwise norms of these derivatives. For the second derivatives of f we have

$$\mathcal{L}_X|\partial^2 f|^2 = e^{BC}e^{DE}\partial_{BD}fX^A\partial_A\partial_{CD}f = \tag{4.50}$$

$$e^{BC}e^{DE}\partial_{BD}f[\partial_{CD}(X^A\partial_A f) - \partial_C(\partial_D X^A)\partial_A f - \partial_C X^A\partial_{AD}f] \tag{4.51}$$

When $\mathcal{L}_X f = 0$, $\mathcal{L}_X|\partial^2 f|^2$ is a quadratic form in ∂f and $\partial^2 f$. The coefficients of terms quadratic in $\partial^2 f$ contain only first-order derivatives of X; these are continuous if g is C^2. Sobolev embedding and multiplication properties permit the estimate of the other terms.

To compute the Lie derivative of derivatives of f of arbitrary order, we can use an inductive argument. For an arbitrary function F on phase space we have

$$\mathcal{L}_X\frac{\partial F}{\partial x^A} = \frac{\partial}{\partial x^A}\mathcal{L}_X F - \frac{\partial F}{\partial x^B}\frac{\partial X^B}{\partial x^A}. \tag{4.52}$$

Induction gives an equality of the form, with symmetrization in $A_1, \ldots, A_k$ and C_ℓ some numbers,

$$\mathcal{L}_X\frac{\partial^k F}{\partial x^{A_1}\ldots\partial x^{A_k}} \equiv \frac{\partial^k(\mathcal{L}_X F)}{\partial x^{A_1}\ldots\partial x^{A_k}} + \sum_{0\le\ell\le k} C_\ell\frac{\partial^{k-\ell}F}{\partial x^{A_1}\ldots\partial x^{A_{k-\ell}}\partial x^B}\frac{\partial^\ell X^B}{\partial x^{A_{k-\ell+1}}\ldots\partial x^{A_k}}. \tag{4.53}$$

In the purely gravitational case we denote $X_x = (p^\alpha)$, $X_p = (-\Gamma^\beta_{\lambda\mu}p^\lambda p^\mu)$. The derivatives of X_x of order greater than 1 are all zero, and the derivatives of X_p are such that

$$\left|\frac{\partial^\ell X_p^B}{\partial x^{A_{k-\ell+1}}\ldots\partial x^{A_k}}\right| \le C|p|^2 \quad \text{if all the } A \text{ indices are spacetime indices}$$

$$\left|\frac{\partial^\ell X_p^B}{\partial x^{A_{k-\ell+1}}\ldots\partial x^{A_k}}\right| \le C|p| \quad \text{if one of the } A \text{ indices is a fiber index} \tag{4.54}$$

$$\left|\frac{\partial^\ell X_p^B}{\partial x^{A_{k-\ell+1}}\ldots\partial x^{A_k}}\right| \le C \quad \text{if two of the A indices are fiber indices.} \tag{4.55}$$

These remarks lead to the following definition, where $\partial^k f$ denotes the set of derivatives of f order $|k|$ and $|.|$ the corresponding Euclidean norm.

Definition 4.7 *The P-energy of order s of the distribution function f at time t is*

$$E_P^{(s)}(t) := \sum_{|k|=1,\ldots s}\int_{\hat{M}_t}(p^0)^{P+k_p}|\partial^k f|^2 N\mu_{\bar{g}}\mu_p, \tag{4.56}$$

The number k_p is the number of derivatives with respect to the $p's$ of order less than or equal to 2 which appear in the derivative of order $|k|, \partial^k f$.

We sketch the proof of the following theorem, a general P, s-energy estimate.

Theorem 4.8 *On a regularly sliced strip V_T of a spacetime (V, g), with*[20] *$g \in \tilde{E}_{s+1}$, the P, s-energy estimate satisfies an inequality*

$$E_P^{(s)}(t) \leq E_P^{(s)}(0) e^{Ct}, \quad t \in [0, T], \tag{4.57}$$

where C depends only on the $\tilde{E}_{s+1}$ bounds of g.

Proof This follows the same lines as the proof of the previous theorem.

The P, s-energy estimate is the integral on the positively oriented manifold $\hat{M}_t$ of the exterior form

$$\sum_{|k|=1,\dots s} (p^0)^{P+k_p} |\partial^k f|^2 i_X \theta, \quad P' = P + 2. \tag{4.58}$$

This definition implies that

$$E_P^{(s)}(t) - E_P^{(s)}(0) = \int_0^t \int_{\hat{M}_\tau} \mathcal{L}_X \left\{ \sum_{|k|=1,\dots s} (p^0)^{P+k_p} |\partial^k f|^2 \right\} \theta \tag{4.59}$$

The general formula applied to $F = f$ with f solution of the Liouville–Vlasov equation leads to an integral inequality

$$E_P^{(s)}(t) - E_P^{(s)}(0) \leq C \int_0^t |E_P^{(s)}(\tau) d\tau \tag{4.60}$$

from which the results follows. □

4.4 *Existence theorem*

As said before, the Cauchy problem for the Liouville–Vlasov equation on an $n+1$-dimensional spacetime (V, g) is the data on the phase space $\mathcal{P}_M$, bundle over an n-dimensional submanifold $\bar{M}$ of V, of a function $\bar{f}$. The following theorem results from the general theory and the estimates of the previous subsection. We leave to the reader the formulation of the relevant functional spaces.

Theorem 4.9 *The Cauchy problem for the Liouville–Vlasov equation on the phase space $\mathcal{P}_{V_T}$ of a regularly sliced strip V_T of an $n+1$-dimensional spacetime (V, g), with data $\bar{f}$ on $\mathcal{P}_{\bar{M}}$, $\bar{M} := M \times \{0\}$ admits one and only one solution f on $\mathcal{P}_{V_T}$.*

Remark that this theorem asserts global existence of the distribution function on the phase space $\mathcal{P}_V$ of a regularly sliced Lorentzian manifold (V, g) – an expected result since the Vlasov equation is linear. It is probably possible to prove the global existence on the tangent bundle of a globally hyperbolic manifold with appropriate hypotheses on the initial manifold and initial data.

[20] See definition in Appendix III.

4.5 *Stress energy tensor of a distribution function*

To apply the local existence theorem proved in Chapter 6 for the Einstein equations with sources, we need the stress energy tensor to belong to a space H_s on slices M_t of the spacetime manifold V.

4.5.1 L^2 *norm on a space slice*

Lemma 4.10 *Let T be the stress energy tensor of a distribution f on the phase space $\mathcal{P}_V$ over an $n+1$-dimensional spacetime (V, g). The tensor T is bounded in L^2 norm on a slice M_t of a regularly sliced strip V_T of (V, g) if the $E_P(t)$ energy of f, $P > n + 5$, is bounded on $\hat{M}_t := \mathcal{P}_{M_t}$.*

Proof The Cauchy–Schwarz inequality implies that the pointwise e norm of the stress energy tensor of a distribution function f is such that, with $|.|$the Euclidean norm

$$|T(x,t)|^2 := |\int_{\mathcal{P}_{x,t}} f(x,t,p)\ p \otimes p \mu_p|^2 \leq A \int_{\mathcal{P}_{x,t}} f^2(x,t,p)\ |p|^P \mu_p \tag{4.61}$$

with

$$A := \int_{\mathcal{P}_{x,t}} |p|^{4-P} \mu_p. \tag{4.62}$$

On the fibre $\mathcal{P}_{x,t}$ the volume element μ_p is equal to the Euclidean volume element $d^{n+1}p$, up to the product by a constant on this fibre, $(\det g)^{\frac{1}{2}}$. Since $|p|$ is the Euclidean norm we have

$$|p| \geq p^0. \tag{4.63}$$

We have assumed that on a fibre

$$p^0 \geq \{N^{-2} g_{ij} p^i p^j + m^2\}^{\frac{1}{2}} \geq m > 0, \tag{4.64}$$

hence

$$|p| \geq m > 0. \tag{4.65}$$

Elementary calculus tells us that the integral A is a finite number as soon as

$$P > n + 5. \tag{4.66}$$

We deduce from (4.64) that over a regular slice M_t there exists a positive constant C

$$p^0 \geq C|p|. \tag{4.67}$$

The L^2 norm of T on the slice M_t therefore satisfies the inequality

$$||T||^2_{L^2(M_t)} \leq AC^{-P} \int_{M_t} \int_{\mathcal{P}_{x,t}} f^2(x,t,p)\ |p^0|^P N \mu_{\bar{g}} \mu_p \equiv AC^{-P} E_P(t). \tag{4.68}$$

□

4.5.2 *H_s norm on a space slice*

The H_s norms of T on M_t are obtained by integrating over M_t the squares of the norms of space derivatives of order up to s. To estimate the H_s norm we bound the pointwise norms of order $|k| \leq s$.

$$|\bar{\partial}^k T(x,t)|^2 = \left| \int_{\mathcal{P}_{x,t}} \bar{\partial}^k f(x,t,p)\ p \otimes p \mu_p \right|^2 . \tag{4.69}$$

The same proof as the one given for the L^2 norm gives the following theorem.

Theorem 4.11 *Let T be the stress energy tensor of a distribution f on the phase space $\mathcal{P}_V$ over an $n+1$-dimensional spacetime (V,g) defined by Equation (5.2). The tensor T is bounded in H_s norm on a slice M_t of a regularly sliced strip V_T of (V,g) if the $E_P^{(s)}(t)$ energy of f, $P > n+5$, is bounded on $\hat{M}_t := \mathcal{P}_{M_t}$.*

5 The Einstein–Vlasov system

The Einstein equations with source a stress energy tensor T read in spacetime dimension $n+1$

$$R_{\alpha\beta} = \rho_{\alpha\beta} \qquad \text{with} \qquad \rho_{\alpha\beta} \equiv T_{\alpha\beta} - \frac{1}{n-1} g_{\alpha\beta} T^{\lambda}_{\lambda}. \tag{5.1}$$

If T is the stress energy tensor of a distribution function f, one has

$$T_{\alpha\beta}(x) := \int_{\mathcal{P}_x} f(x,p) p_\alpha p_\beta \mu_p. \tag{5.2}$$

The Einstein equations are coupled with the Liouville–Vlasov equation for f

$$\mathcal{L}_X f \equiv p^\alpha \frac{\partial f}{\partial x^\alpha} + Q^\alpha \frac{\partial f}{\partial p^\alpha} = 0, \quad Q^\alpha := -\Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu. \tag{5.3}$$

5.1 *Constraints*

An initial data set for the Einstein–Vlasov system is a quadruplet $(M, \bar{g}, K, \bar{f})$, with $(M, \bar{g}, K)$ Einstein's initial data set and $\bar{f}$ a function on the bundle $\mathcal{P}_M$ corresponding to the kind of particles we consider. As always the data must satisfy the Hamiltonian constraint

$$R(\bar{g}) - |K|^2_{\bar{g}} + (tr_{\bar{g}} K)^2 = 2N^2 T^{00}, \tag{5.4}$$

and the momentum constraint

$$\bar{\nabla}.K - \bar{\nabla} tr K = N T^{0i}. \tag{5.5}$$

If $f(x,p)$ is symmetric in the space component of p,

$$f(x,p^0,p^i) = f(x,p^0,-\bar{p}), \quad \bar{p} = (p^I) \tag{5.6}$$

then the current J source of the momentum constraint vanishes:

$$T^{0i} := \int_{\mathcal{P}_x} f(.,p^0,\bar{p})p^0 p^i \omega = 0, \tag{5.7}$$

therefore, the conformally formulated constraints with kinetic source decouple if the initial manifold has constant mean extrinsic curvature.

5.2 *Cauchy problem for the Einstein equations*

The Einstein equations in wave gauge with source a distribution function f are a quasilinear hyperquasilinear system of wave equations on f with source the stress energy tensor of f. They read

$$R^{(h)}_{\alpha\beta} \equiv -\frac{1}{2} g^{\lambda\mu}\partial^2_{\lambda\mu} g_{\alpha\beta} + H_{\alpha\beta}(g)(\partial g, \partial g) = \rho_{\alpha\beta}. \tag{5.8}$$

We have seen in Chapter 6 how to deduce from an initial data set $(M, \bar{g}, K)$ for the geometric Einstein equations initial data $g(0,.)$, $\frac{\partial g}{\partial t}(0,.)$ on a submanifold $M_0 := M \times \{0\}$ embedded in $M \times R$ for the Einstein equations in wave gauge. These initial data are in[21] $\tilde{H}_s \times H_{s-1}$ if $(\bar{g}, K) \in \tilde{H}_s \times H_{s-1}$, $s > \frac{n}{2} + 1$. We have shown in Appendix III that such a Cauchy problem for equations like the Einstein equations in wave gauge with source a given stress energy tensor $T \in E_{s-1}(L)$ on a strip $V_L := [0, L] \times M$ has a solution $g \in \tilde{E}_s$ on a strip $V_\ell := [0, \ell] \times M$ regularly sliced by g if $\ell \leq L$ is small enough.

5.3 *Cauchy problem for the coupled system*

5.3.1 *Local existence theorem*

The following theorem is to be proved by iteration or use of the contracting map principle.

Theorem 5.1 *The Cauchy problem with data satisfying the constraints, $(\bar{g}, K) \in \tilde{H}_s \times H_{s-1}$ on M, $\bar{f} \in H_{s-1}(\mathcal{P}_M)$, $s > \frac{n}{2} + 2$, for the Einstein–Vlasov system admits a solution (g, f) on $V_\ell \times \mathcal{P}_V$, with $V_\ell := M \times [0, \ell])$, development of the initial data set $(M, \bar{g}, K)$, $M_0 = M \times \{0\}$, $g \in \tilde{E}_s(\ell), f \in \hat{E}_{s-1}(\ell)$.*

Proof Let the pair (g_1, f_1) be a Lorentzian metric and a distribution function, respectively, on a strip V_L and the bundle $\mathcal{P}_{V_L}$, $g_1 \in \tilde{E}_s(L)$, $f_1 \in \hat{E}_{s-1,P}(L)$. Consider the integro-differential system with unknowns g_2 and f_2

$$R^{(h)}_{\alpha\beta}(g_2) = \rho_{\alpha\beta}(f_1), \quad \mathcal{L}_{X_1} f_2 = 0 \tag{5.9}$$

where $R^{(h)}_{\alpha\beta}(g_2)$ is the Ricci tensor of g_2 in wave gauge and X_1 is the tangent vector to the geodesic flow of g_1. There is a strip V_ℓ, $\ell \leq L$ such that the system (5.1), (5.2) has one and only one solution $g_2 \in \tilde{E}_s(\ell)$, $f_2 \in \hat{E}_{s-1}(\ell)$ taking the given Cauchy data. The proof that the mapping is contracting, for small enough ℓ, is

[21] Recall that $\bar{\mathrm{g}} \in \tilde{\mathrm{H}}_s$ if $\bar{\mathrm{g}} \in \bar{\mathrm{C}}^0$ and $\partial\, \bar{\mathrm{g}} \in\ \mathrm{H}_{s-1}$.

straightforward, using the energy inequalities for quasilinear wave equations and the energy estimates of Section 4.3. The fixed point (g, f) satisfies the Vlasov equation, hence f gives a stress energy tensor T with vanishing divergence in the metric g. The metric g satisfies the Einstein equations in wave gauge with source T. Standard arguments show that the full Einstein–Vlasov equations are satisfied if the initial data satisfy the constraints. □

5.3.2 *Global existence theorems*

The collisionless kinetic theory, having no global problems of its own, one may hope to extend to the Einstein–Vlasov system global results obtained for the vacuum Einstein equations. Such results are already hard to prove (see Chapters 15 and 16). Global existence theorems, or proofs of the cosmic censorship conjectures, for the Einstein–Vlasov system have been obtained only in the presence of an isometry group. The first global existence theorem, for small initial data, was proved by Rein and Rendall in the case of spherical symmetry[22]. A recent paper by Dafermos and Rendall[23] proves cosmic censorship in the case of surface symmetric compactly supported initial data. References to other global results concerning the Einstein–Vlasov system can be found in that paper and also in a review article by Andréasson[24].

6 The Einstein–Maxwell–Vlasov system

When, in addition to gravitation, the particles are subjected to an electromagnetic field, a 2-form F, the component in the fibre $\mathcal{P}_V$ of the vector X tangent to the particles trajectories is the sum of the gravitational component $Q^\alpha := -\Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu$ and the electromagnetic force, $\Phi^\alpha := e p^\beta F_\beta{}^\alpha$ if all particles have the same electric charge e. The orthogonality of p and Φ permits the proof of the extension to the Maxwell–Vlasov equation

$$\mathcal{L}_X f := p^\alpha \frac{\partial f}{\partial x^\alpha} + (Q^\alpha + \Phi^\alpha) \frac{\partial f}{\partial p^\alpha} = 0 \tag{6.1}$$

of the results obtained for the Liouville–Vlasov equations.

The Maxwell–Vlasov equation must be coupled with an Einstein–Maxwell system with source f.

The Maxwell equations satisfied by an electromagnetic 2-form F, with source an electric current J, are

$$dF = 0, \quad \delta F = J, \quad \text{i.e.} \quad \nabla_\alpha F^{\alpha\beta} = J^\beta. \tag{6.2}$$

In the Maxwell–Vlasov system J is the average electromagnetic current generated by the charged particles; that if all particles have the same charge e,

$$J^\alpha = e \int_{\mathcal{P}_x} f(x, p) p^\alpha \omega_x. \tag{6.3}$$

[22] Rein, D. and Rendall, A. (1992) *Commun. Math. Phys.*, **150**, 561–83.
[23] Dafermos, M. and Rendall, A. (2007) arXiv gr-qc 0610075v1.
[24] Andréasson H. 2005 arXiv gr-qc/0502091 v1.

Remark 6.1 *If there are N species of charged particles with masses m_I, charges e_I, and distribution functions f_{m_I}, the mean electric current they generate is*

$$J^\alpha = \sum_I e_I \int_{\mathcal{P}_{m_I,x}} f_{m_I}(x,p) p^\alpha \omega_{m_{I,p}}. \tag{6.4}$$

The Maxwell–Vlasov system in Special Relativity has been studied extensively[25].

A **Vlasov–Maxwell–Einstein system** is the set of equations (6.1), (6.2) with J given by (6.3), and the Einstein equations with source the sum of the stress energy tensor τ of F and of the second moment of the distribution f, i.e.

$$S^{\alpha\beta} = T^{\alpha\beta} \equiv F^\alpha{}_\lambda F^{\beta\lambda} - \frac{1}{4} g^{\alpha\beta} F_{\lambda\mu} F^{\lambda\mu} + \int_{\mathcal{P}_x} f(x,p) p^\alpha p^\beta \omega_p. \tag{6.5}$$

The following theorem makes the equations coherent.

Theorem 6.2

1. *If f satisfies the Maxwell–Vlasov equation (6.1) then the current J has a zero divergence.*
2. *If, in addition, F satisfies the Maxwell equations (6.2), then the stress energy tensor T is divergence free.*

Proof The proofs can be made as in Theorem 3.5 by choosing coordinates such that the Christoffel symbols vanish at the considered point. We give them in the more general case of Remark (2.6).

1. The Maxwell–Vlasov equation implies that

$$\frac{\partial J^\alpha}{\partial x^\alpha} = \sum_I e_I \int_{\mathcal{P}_{m_I,x}} \Phi^\alpha \frac{\partial f_{m_I}(x,p)}{\partial p^\alpha} \omega_{m_{I,p}}. \tag{6.6}$$

We have, using integration by parts and the property $\partial \Phi^\alpha / \partial p^\alpha = 0$,

$$\int_{\mathcal{P}_{m_I,x}} \Phi^\alpha \frac{\partial f_{m_I}(x,p)}{\partial p^\alpha} \omega_{m_{I,p}} = - \int_{\mathcal{P}_{m_I,x}} f_{m_I}(x,p) \frac{\partial \Phi^\alpha}{\partial p^\alpha} \omega_{m_{I,p}} = 0. \tag{6.7}$$

2. We have seen in Chapter 3 that the divergence of the Maxwell tensor is if F satisfies the Maxwell equations

$$\nabla_\alpha \tau^{\alpha\beta} = J^\lambda F_\lambda{}^\beta \tag{6.8}$$

We have, for any function $f(x,p)$ at the considered point

$$\nabla_\alpha \int_{\mathcal{P}_{m_I,x}} f(x,p) p^\alpha p^\beta \omega_{m_{I,p}} = \int_{\mathcal{P}_{m_I,x}} \frac{\partial f(x,p)}{\partial x^\alpha} p^\alpha p^\beta \omega_{m_{I,p}}. \tag{6.9}$$

[25] In particular by Glassey and Strauss.

Hence if f_{m_I} satisfies the Maxwell–Vlasov equation

$$\nabla_\alpha \int_{\mathcal{P}_{m_I,x}} f_{m_I}(x,p)p^\alpha p^\beta \omega_{m_I,p} = -\int_{\mathcal{P}_{m_I,x}} \Phi^\alpha \frac{\partial f_{m_I}(x,p)}{\partial p^\alpha} p^\beta \omega_{m_I,p}. \tag{6.10}$$

The calculus identity

$$\Phi^\alpha \frac{\partial f_{m_I}(x,p)}{\partial p^\alpha} p^\beta \equiv \frac{\partial}{\partial p^\alpha}[\Phi^\alpha f_{m_I}(x,p)p^\beta] - \frac{\partial \Phi^\alpha}{\partial p^\alpha} f_{m_I} p^\beta - \Phi^\beta f_{m_I} \tag{6.11}$$

integration by parts and the property $\partial\Phi^\alpha/\partial p^\alpha = 0$ show that

$$\nabla_\alpha \sum_I e_I \int_{\mathcal{P}_{m_I,x}} f_{m_I}(x,p)p^\alpha p^\beta \omega_{m_I,p} = \sum_I e_I \int_{\mathcal{P}_{m_I,x}} \Phi^\beta f_{m_I} \omega_{m_I,p}. \tag{6.12}$$

The definitions of Φ and J show the equality, up to the sign of the right-hand sides, of (6.8) and (6.12); hence the vanishing of the divergence of the tensor T given by 6.5. □

7 Boltzmann equation. Definitions

When particles undergo collisions their trajectories in phase space are no longer connected integral curves of the vector field X since their momenta undergo a jump with the crossing of another trajectory. Consequently the derivative of the distribution function f along X is no longer zero. In the Boltzmann model this derivative is equal to the so-called collision operator $\mathcal{I}f$:

$$\mathcal{L}_X f = \mathcal{I}f \tag{7.1}$$

$\mathcal{I}f$ is an integral operator which is to be interpreted as being linked with the probability that two particles of momentum respectively p' and q' collide at x and give after the shock two particles, one of them with momentum p and the other one with momentum q. One says that the shock is elastic when the following law[26] of conservation of momentum holds:

$$p' + q' = p + q \tag{7.2}$$

This equation defines a submanifold Σ in the fibre $(\times\mathcal{P}_x)^4$. For fixed p, q, one denotes by Σ_{pq} the submanifold of $(\times\mathcal{P}_x)^2$ defined by (7.2) The volume element ξ' (Leray form) in Σ_{pq} is such that

$$\xi' \wedge (\bigwedge_\alpha (d(p'^\alpha + q'^\alpha)) = \omega_{p'} \wedge \omega_{q'}. \tag{7.3}$$

The collision operator is

$$(\mathcal{I}f)(x,p) \equiv \int_{\mathcal{P}_x(q)} \int_{\Sigma_{pq}} [f(x,p')f(x,q') - f(x,p)f(x,q)]A(x,p,q,p',q')\xi' \wedge \omega_q. \tag{7.4}$$

[26] This law supposes no change in internal properties of the particles.

The function $A(x,p,q,p',q')$ is called the shock cross-section. It is a phenomenological quantity. No explicit expression is known for it in Relativity. A generally admitted property is the reversibility of elastic shocks, namely:

$$A(x,p,q,p',q') = A(x,p',q',p,q). \tag{7.5}$$

Lemma 7.1 *When the particles all have the same, non-zero, proper mass m the integral on Σ_{pq} can be written by using a formula of the type, with θ, φ canonical angular parameters on the sphere S^2 (case $n+1=4$)*

$$A(x,p,q,p',q')\xi' = S(x,p,q,\theta,\varphi) sin\theta d\theta \wedge d\varphi.$$

Proof Introduce at the point x, for a given pair p,q of timelike vectors, an orthonormal Lorentz frame with time axis e_0 in the direction of $p+q$. Set $p+q = 2\lambda e_0$, $2\lambda = \{-g(p+q)\}^{1/2}$. Then also $p'+q' = 2\lambda e_0$, and since p,q,p',q' are timelike vectors with length m,

$$p^\alpha q_\alpha = p'^\alpha q'_\alpha = m^2 - 2\lambda^2 \leq -m^2.$$

In such a frame $p^0 + q^0 = p'^0 + q'^0 = 2\lambda$ while $p^i + q^i = p'^i + q'^i = 0$; hence $\Sigma(p^i)^2 = \Sigma(q^i)^2$ and therefore $p^0 = q^0 = \lambda$ while $\Sigma(p^i)^2 = \lambda^2 - m^2 =: \alpha^2$. The same properties hold for the prime variables. In this frame Σ_{pq} is represented by a 2-sphere, $\Sigma(p'^i)^2 = \alpha^2$ of radius α in the plane $p'^0 = \lambda$. It holds that:

$$\alpha = \frac{1}{2}(-p^\alpha q_\alpha - m^2)^{1/2} = \frac{1}{2} g(p-q, p-q)^{1/2}.$$

Take a vector parallel to $p-q$ (which is orthogonal to $p+q = 2\lambda e_0$) as the axis for polar coordinates θ, φ on Σ_{pq}, denoting by θ the angle of the space vectors $p-q$ and $p'-q'$; the definition of ξ' gives

$$\xi' = (2\lambda)^{-1} \alpha sin\theta d\theta \wedge d\varphi.$$

The given relation holds with

$$S(p,q,\theta,\varphi) = A(x,p,q,p',q')(2\lambda)^{-1}\alpha \tag{7.6}$$

where on the sphere Σ_{pq} both p' and q' are given by the θ, φ dependent 4-vectors

$$p'_{\Sigma_{pq}} = q'_{\Sigma_{pq}} = \left\{ -\frac{1}{2} g(p+q)^{1/2}, \right.$$
$$\left. \frac{1}{2} g(p-q, p-q)^{1/2} (cos\theta, sin\theta cos\varphi, sin\theta sin\varphi) \right\}. \tag{7.7}$$

□

8 Moments and conservation laws

The moments of a distribution function f have been defined in Section 2 by integrals over a fibre in phase space. The moment of order n is

$$T^{\alpha_1 \dots \alpha_n}(x) =: \int_{\mathcal{P}_x} p^{\alpha_1} \dots p^{\alpha_n} f(x,p) \omega_p. \tag{8.1}$$

It satisfies the identity

$$\nabla_\alpha T^{\alpha_2 \dots \alpha_n} \equiv \int_{\mathcal{P}_x} p^{\alpha_2} \dots p^{\alpha_n} \mathcal{L}_X f(x,p) \omega_p. \tag{8.2}$$

The right-hand side is zero if f satisfies the Liouville–Vlasov equation.

We have interpreted the first and second moments as giving respectively the proper rest mass energy momentum vector and the stress energy tensor of the macroscopic matter corresponding to the distribution function f. We prove a lemma, important for coherence with the Einstein equations, when f satisfies a Boltzmann equation with reversible collisions.

Lemma 8.1 *The first and second moments of a distribution f satisfying a Boltzmann equation on a spacetime (V, g) have a zero divergence in the spacetime metric if the collisions are reversible.*

Proof The first moment of f is the vector field on V, interpreted as proper rest mass energy momentum, given by

$$P^\alpha := \int_{\mathcal{P}_x} p^\alpha f(x,p) \omega_p. \tag{8.3}$$

Equation (8.2) reads in this case

$$\nabla_\alpha P^\alpha \equiv \int_{\mathcal{P}_x} \mathcal{L}_X f(x,p) \omega_p = \int_{\mathcal{P}_x} (\mathcal{I} f)(x,p) \omega_p. \tag{8.4}$$

The second moment of f, interpreted as the stress energy tensor, is defined by

$$T^{\alpha\beta} := \int_{\mathcal{P}_x} p^\alpha p^\beta f(x,p) \omega_p. \tag{8.5}$$

For f satisfying a Boltzmann equation we have

$$\nabla_\alpha T^{\alpha\beta} \equiv \int_{\mathcal{P}_x} p^\beta (\mathcal{I} f)(x,p) \omega_p. \tag{8.6}$$

Standard calculus, which we leave to the reader as an exercise, shows that, for reversible collisions, Equations (8.4) and (8.6) have a zero right-hand side.

In the case of particles with non-zero rest mass we have set $P^\alpha = r u^\alpha$, with u the unit flow vector of the macroscopic matter corresponding to the distribution function f. The Equation (8.4) is identical to the matter conservation law found for fluids,

$$\nabla_\alpha (r u^\alpha) = 0. \tag{8.7}$$

Equation (8.6)

$$\nabla_\alpha T^{\alpha\beta} = 0 \tag{8.8}$$

is to be satisfied by all stress energy tensors in General Relativity. □

The higher moments do not satisfy conservation laws, but form an infinite series; see Section 11.

9 Einstein–Boltzmann system

The Einstein equations with source the stress energy tensor of a distribution function f satisfying a Boltzmann equation are a coherent system if the collision operator is such that the stress energy tensor is conservative. We have proved in Lemma 8.1 that this property holds if the collisions are reversible.

It has been proved[27] that the local Cauchy problem, with initial data set $(M, \bar{g}, K, \bar{f})$ is well posed for such a Einstein–Boltzmann system.

The same kind of results hold for an Einstein–Maxwell–Boltzmann system obtained in replacing the Liouville–Vlasov operator by the Liouville–Maxwell–Vlasov operator.

10 Thermodynamics

One of the main interests of the relativistic kinetic theory is in the possibility of obtaining the laws of thermodynamics in a relativistic context. The problem is already difficult in Special Relativity. New problems arise in General Relativity due to the non-existence of an equilibrium distribution function in a non-stationary universe.

10.1 *Entropy and H theorem*

The entropy density of a distribution f is the function $-k(f \log f)(x,p)$ on phase space, where k is the Boltzmann constant[28]. The entropy flux on spacetime is the future directed timelike vector[29]

$$H^\alpha(x) := -k \int_{\mathcal{P}_x} p^\alpha (f \log f)(x,p) \omega_p \tag{10.1}$$

The following theorem is the relativistic formulation of a theorem well-known in non-relativistic thermodynamics.

Theorem 10.1 *(H theorem) If collisions are reversible and satisfy the symmetry property (7.5) the entropy flux H^α is such that*

$$\nabla_\alpha H^\alpha \geq 0. \tag{10.2}$$

Proof The divergence of H is found to be, by a computation similar to a previous one,

$$(\nabla_\alpha H^\alpha)(x) \equiv -k \int_{\mathcal{P}_x} (\mathcal{L}_X [f \log f])(x,p) \omega_p \equiv -k \int_{\mathcal{P}_x} (\mathcal{L}_X f[\log f + 1])(x,p) \omega_p \tag{10.3}$$

[27] Bancel, D. (1970) *Ann. Inst. Poincaré* (Boltzmann equation); Bancel, D. and Choquet-Bruhat, Y. (1973) *Commun. Maths. Phys.* (coupled system).

[28] $k = 1.3806568 \times 10^{-23}$ Joule per Kelvin is the quotient of the perfect gas constant and Avogadro's number.

[29] Under the assumption due to the physical definition of f one has $0 < f < 1$; i.e. $\log f < 0$.

Using the Boltzmann equation and the reversibility of collisions which implies that

$$\int_{\mathcal{P}_x(q)} \int_{\Sigma_{pq}} [f(x,p')f(x,q') - f(x,p)f(x,q)]A(x,p,q,p',q')\xi' \wedge \omega_q \wedge \omega_p = 0 \tag{10.4}$$

we find that

$$(\nabla_\alpha H^\alpha)(x) = -k \int_{\mathcal{P}_x(q)} \int_{\Sigma_{pq}} [f(x,p')f(x,q') - f(x,p)f(x,q)](logf)(x,p)A(x,p,q,p',q')\xi' \wedge \omega_q \wedge \omega_p. \tag{10.5}$$

Making moreover the natural assumption of symmetry:

$$A(x,p,q,p',q') = A(x,q,p,p',q'), \tag{10.6}$$

one rewrites the above divergence as follows (the dependence on x has been made implicit for shortness):

$$\nabla_\alpha H^\alpha = -\frac{k}{4} \int_{\mathcal{P}_x(q)} \int_{\Sigma_{pq}} [f(p')f(q') - f(p)f(q)]log\frac{f(p)f(q)}{f(p')f(q')}A(p,q,p',q')\xi' \wedge \omega_q \wedge \omega_p. \tag{10.7}$$

which is non-negative[30] when A is non-negative. □

The gas is said to be in thermal equilibrium if $\nabla_\alpha H^\alpha = 0$. A general relativistic gas undergoing collisions will not in general attain in a finite time thermal equilibrium (see next section).

We can make a simple link between the microscopic and macroscopic properties of the gas when the distribution f is isotropic in phase space, namely if there exists a timelike vector v on spacetime such that f depends only on $v_\alpha p^\alpha$. We have seen (Section 2) that the macroscopic gas is then a perfect fluid with momentum flow P collinear with V. The same type of proof gives the following lemma.

Lemma 10.2 *Assume there exists on spacetime a timelike vector V such that the function f is symmetric in the spaces orthogonal to V, i.e. in each fibre,*

$$f(x,q) = f(x,p) \quad \textit{if} \;\; V_\alpha p^\alpha = V_\alpha q^\alpha \;\; \textit{and} \quad \bar{p} = -\bar{q} \tag{10.8}$$

where $\bar{p}$ and $\bar{q}$ denote respectively the projections of p and q on the subspace orthogonal to V. Then the first moment $P = rU$ as well as the entropy vector H are collinear with V.

[30] Because $(a-b)(\log a - \log b) \geq 0$.

Lemma 10.3 *When the vectors H and P are collinear one defines a positive specific scalar entropy density S by setting*

$$H^\alpha = SP^\alpha. \tag{10.9}$$

This entropy density S satisfies on spacetime the inequality

$$P^\alpha \partial_\alpha S \geq 0. \tag{10.10}$$

Proof The inequality $\nabla_\alpha H^\alpha \geq 0$ and the conservation law $\nabla_\alpha P^\alpha = 0$. □

When f does not have the property (10.8) the entropy and matter flux are not collinear.

In all cases integrating (10.2) on a spacelike slice V_T, with compact space or appropriate boundary conditions at spacelike infinity, we find that:

$$\int_{M_T} H^0 N \mu_{\bar{g}} \geq \int_{M_0} H^0 N \mu_{\bar{g}}. \tag{10.11}$$

In an expanding universe, for instance a Robertson–Walker universe, where H^α depends only on t, we define a total entropy of space by integrating on space the component H^0. Denoting this total entropy on space by Σ we find that:

$$\frac{d(R^3\Sigma)}{dt} \geq 0. \tag{10.12}$$

where R is the Robertson–Walker radius. We see that a decrease of Σ is linked with an increase of R. *It is the expansion of the universe which permits its ever-increasing organization; that is, the decrease of its entropy from an initial anisotropy of f.*

10.2 *Maxwell–Jüttner equilibrium distribution*

A gas is considered as in thermal equilibrium if its entropy is conserved in the following sense:

$$\nabla_\alpha H^\alpha = 0. \tag{10.13}$$

The formula (10.3) shows that a sufficient condition for this equality to hold is that the distribution function f be conserved along the trajectories of X in phase space, i.e. $\mathcal{L}_X f = 0$. This condition is in a Boltzmann framework,

$$\mathcal{I}(f) = 0. \tag{10.14}$$

A sufficient condition for this equality to hold is that at each point x of spacetime the function f is such that:

$$f(p')f(q') - f(p)f(q) = 0, \quad \text{if} \quad p + q = p' + q'. \tag{10.15}$$

This condition is also necessary if $A(p, q, p', q')$ is strictly positive. It can also be proved to hold under weaker assumptions (Marle, 1969, p. 94). It is immediate to check that a solution of the above functional equation is:

$$f(x, p) = a(x) exp(b_\alpha(x) p^\alpha) \tag{10.16}$$

with a a positive scalar function and b a covariant vector on spacetime. It can be proved that all continuous solutions of (10.15) are of this form. If a is non-negative the same is true of f; if, moreover, b is future timelike then $b_\alpha(x)p^\alpha < 0$ and $f(x,.)$ is integrable on $\mathcal{P}_x$. Such functions are called **Maxwell–Jüttner distributions**.

For such a Maxwell–Jüttner distribution f we find by straightforward computation that:

$$\mathcal{L}_X f \equiv e^{b_\lambda p^\lambda} \left\{ p^\alpha \frac{\partial a}{\partial x^\alpha} + \frac{1}{2} p^\alpha p^\beta (\nabla_\alpha b_\beta + \nabla_\beta b_\alpha) \right\} \tag{10.17}$$

A function f given by (10.16) will satisfy $\mathcal{L}_X f = 0$ on the phase space if and only if on spacetime it holds that $\frac{\partial a}{\partial x^\alpha} = 0$ and $\nabla_\alpha b_\beta + \nabla_\beta b_\alpha = 0$ that is iff a is a constant and b is a Killing vector field: the spacetime must therefore be stationary to admit a thermal equilibrium distribution function, an unsurprising result!

We deduce from Theorem 2.4 that the rest mass momentum vector and the stress energy tensor on spacetime associated with a Maxwell Jüttner distribution with a timelike vector b are those of a perfect fluid with flow vector collinear with b. The same proof as in Theorem 2.4 gives also that the entropy vector H is collinear with b, hence collinear with u. The integrals giving the moments can be expressed through the modified Bessel functions of the second kind and index n (see Pichon, 1967; Marle, 1969). The scalar $\lambda := (-b^\alpha b_\alpha)^{\frac{1}{2}}$ is interpreted as the inverse of the product of the absolute temperature by the Boltzmann constant:

$$\lambda = (kT)^{-1}. \tag{10.18}$$

10.3 *Dissipative fluids*

In classical mechanics the Navier–Stokes equations can be obtained from a distribution function[31] which is a first-order perturbation of the Maxwell equilibrium distribution, either by the Chapman–Enskog method or by the Grad polynomials expansion method. Both methods have been extended by Marle (see also Stewart) who found the stress energy tensor corresponding to a first-order perturbation of the Maxwell–Jüttner distribution. The corresponding, very complex, system of equations have been shown to be of hyperbolic type only in some simplified cases. The fact that thermal equilibrium is incompatible with non-stationary spacetimes make the validity of such equations doubtful in cosmology.

11 Extended thermodynamics

Contributed by Tommaso Ruggeri

A pioneer of Relativistic Thermodynamics was Carl Eckart, who as early as 1940 established the thermodynamics of irreversible processes, a theory now

[31] Recent work by F. Golse and L. Saint Raymond links the Leray global weak solutions of the Navier–Stokes equations with those of P. L. Lions for the Boltzmann equation.

universally known by the acronym TIP. Eckart's theory is an important step away from equilibria toward non-equilibrium processes. It provides a counterpart of the Navier–Stokes equations for the deviatoric stress and a generalization of Fourier's law of heat conduction. The latter permits a heat flux to be generated by an acceleration, or a temperature gradient to be equilibrated by a gravitational field. But Eckart's theories have one drawback: they lead to parabolic equations for the temperature and velocity and thus predict infinite pulse speeds in contrast with Relativity.

As is well known, the pioneering papers of Müller[32] and Israel[33] were the first attempts to obtain a causal relativistic phenomenological theory that gives a system of equations of hyperbolic type such that wave velocities are finite in agreement with the relativity principle. This approach was based substantially on the idea of modifying the Gibbs relation in non-equilibrium. This procedure was fundamental for a long time because of its simplicity. However, a more refined analysis showed that several degrees of freedom remain and as a consequence some assumptions do not seem completely justified from a "rational" point of view. Moreover, the equations so obtained encountered several mathematical inconveniences. For the former reason a revisited theory was developed in 1986 by Liu, Müller and Ruggeri[34] that requires only very few natural assumptions and uses only universal principles.

11.1 *The phenomenological* 14 *moments theory*

The first paper of the new Extended Thermodynamics (ET) was purely phenomenological also if the basic structure of the equations was taken from the first 14 moments of the kinetic theory.

The objective of extended thermodynamics of relativistic fluids is the determination of the 14 fields $T^{\alpha}(x)$ – particle number, particle flux – and $T^{\alpha\beta}(x)$ – stress, energy–momentum tensor.

To determine the 14 state variables one needs the field equations: the conservation laws of particle number and of energy–momentum

$$\nabla_{\alpha} T^{\alpha} = 0, \qquad \nabla_{\alpha} T^{\alpha\beta} = 0 \tag{11.1}$$

and the (extended) balance laws of fluxes

$$\nabla_{\alpha} T^{\alpha\beta\mu} = I^{\beta\mu}. \tag{11.2}$$

It is assumed that $T^{\alpha\beta}, T^{\alpha\beta\mu}$, and $I^{\beta\mu}$ are completely symmetric tensors and moreover that (according to the kinetic theory):

$$I^{\alpha}{}_{\alpha} = 0, \qquad T^{\alpha\beta}{}_{\beta} = m^2 c^2 T^{\alpha}. \tag{11.3}$$

[32] Müller, I. (1966) *PhD Thesis*, Aachen.
[33] Israel, W. (1976) *Ann. Phys.*, **100**.
[34] I-Shih Liu, I., Müller, T. and Ruggeri, T. (1986) *Ann. Phys.*, **169**(1), 191–219.

As is usual in the continuum approach, the field equations contain unknowns that are not in the list of state variables. In such a case, the tensors $T^{\alpha\beta\mu}$ and $I^{\alpha\beta}$ and therefore constitutive equations are requested. The ET assumes that the constitutive relations are in a local form, i.e.:

$$T^{\alpha\beta\mu} \equiv \hat{T}^{\alpha\beta\mu}(T^{\nu}, T^{\gamma\rho}), \qquad I^{\alpha\beta} \equiv \hat{I}^{\alpha\beta}(T^{\nu}, T^{\gamma\rho}). \tag{11.4}$$

If the functions (11.4) are known, the field equations become a full set of quasilinear first-order partial differential equations, the solutions of which are called "*thermodynamic processes*". Now the problem becomes a constitutive problem, i.e. the restrictions of the acceptable constitutive equations that are compatible with universal physical principles. The universal principles adopted in ET are:

1. **Entropy principle**: there exists a 4-entropy vector H^{α} that is a constitutive quantity

$$H^{\alpha} \equiv \hat{H}^{\alpha}(T^{\nu}, T^{\gamma\rho}) \tag{11.5}$$

such that

$$\nabla_{\alpha} H^{\alpha} \geq 0 \tag{11.6}$$

for all thermodynamic processes.

2. **Relativity principle**: The field equations have the same form in all frames. This statement implies that the constitutive functions are invariant under a change of frame.
3. **Convexity and hyperbolicity**: There exists a timelike vector ξ_{α} for which the function $S = H^{\alpha}\xi_{\alpha}$ is a strictly concave function.

This condition implies that the differential system is a symmetric hyperbolic one and therefore we have not only causality but also well-posed (local in time) Cauchy problem[35]. Moreover this condition is justified also from a physical point of view, because it guarantees that the density of entropy S has a maximum in equilibrium and therefore we obtain also the thermodynamical stability.

Even if the problem is conceptually simple, it is very hard to obtain the constitutive functions (11.4) and (11.5) that satisfy the universal principles. The key to solving the problem consists in an appropriate technique for exploiting the entropy principle that holds for a generic quasilinear first-order system of balance laws. This mathematical procedure is based on the choice of the privileged main field $\mathbf{u}'$ introduced by Ruggeri and Strumia (1981) for which the original system becomes symmetric hyperbolic. For processes not so far from an equilibrium state the results are the following (we use now the usual physical

[35] Ruggeri, T. and Strumia, A. (1981) *Ann. Inst. H. Poincaré*, **34A**, 65–84.

variables representation):

$$
\begin{aligned}
T^{\alpha} &= ru^{\alpha},\\
T^{\alpha\beta} &= t^{<\alpha\beta>} + (p+\pi)h^{\alpha\beta} + \frac{1}{c^2}(u^{\alpha}q^{\beta} + u^{\beta}q^{\alpha}) + \frac{e}{c^2}u^{\alpha}u^{\beta},\\
T^{\alpha\beta\gamma} &= (C_1 + C_2\pi)u^{\alpha}u^{\beta}u^{\gamma} + \frac{c^2}{6}(r - C_1 - C_2\pi)\left(g^{\alpha\beta}u^{\gamma} + g^{\beta\gamma}u^{\alpha} + g^{\alpha\gamma}u^{\beta}\right)\\
&\quad + C_3\left(g^{\alpha\beta}q^{\gamma} + g^{\beta\gamma}q^{\alpha} + g^{\alpha\gamma}q^{\beta}\right) + C_4\left(t^{<\alpha\beta>}u^{\gamma} + t^{<\beta\gamma>}u^{\gamma} + t^{<\alpha\gamma>}u^{\beta}\right)\\
&\quad - \frac{6}{c^2}C_5\left(u^{\alpha}u^{\beta}q^{\gamma} + u^{\beta}u^{\gamma}q^{\alpha} + u^{\alpha}u^{\gamma}q^{\beta}\right),\\
I^{\alpha\beta} &= B_1\pi\left(g^{\alpha\beta} - \frac{4}{c^2}u^{\alpha}u^{\beta}\right) + B_2 t^{<\alpha\beta>} + \frac{1}{c^2}B_3\left(q^{\alpha}u^{\beta} + q^{\beta}u^{\alpha}\right).
\end{aligned}
\tag{11.7}
$$

where $r, t^{<\alpha\beta>}, p, q^{\alpha}, e, \pi$ and u^{α} are respectively the particle density, the stress deviator, the pressure, the heat flux, the energy density, the dynamical pressure and the 4-velocity. The coefficients C_i $(i = 1,2,3,4,5)$ and B_j $(j = 1,2,3)$ are explicit functions of the temperature and the density. The interesting result is that the closure obtained using only macroscopic arguments coincides with that obtained by Marle through some kinetic considerations. For the details the interested reader can see the book of Müller and Ruggeri (see below).

Remark. As was observed in Ruggeri (1987)[36] the point of view of the ET changes the meaning also of the Einstein equations. In fact in the classical approach in which the balance laws are the usual five, only one component (for example the internal energy) in the energy–momentum tensor is considered as the field variable and the remaining ones are given by constitutive equations. In ET it is assumed that all the components of the energy–momentum tensor are field variables and then the Einstein equations become a "universal" equation, independent of the constitution of the material.

As the system is symmetric hyperbolic due to the concavity of the entropy density if the fluxes and the production are smooth enough, in a suitable convex open set $\mathfrak{D}$, it is well known that the system has a unique local (in time) smooth solution for smooth initial data.

However, for general symmetric hyperbolic systems, even for arbitrarily small and smooth initial data, there is no global continuation for these smooth solutions, which may develop singularities, shocks or blow up at finite time.

On the other hand, in many physical examples of systems of balance laws, thanks to the interplay between the source term and the hyperbolicity there exist global smooth solutions for a suitable set of initial data. Roughly speaking, for such a system the relaxation term induces a dissipative effect. This effect

[36] Ruggeri, T. (1987) *Corso CIME Noto*, Lecture Notes in Mathematics 1385 (eds. M. Anile and Y. Choquet-Bruhat), pp. 269–77, Springer-Verlag, Berlin.)

then competes with the hyperbolicity. If the dissipation is sufficiently strong to dominate the hyperbolicity, the system is *dissipative*, and we expect that the classical solution exists for all time and converges to a constant state. Otherwise, if the dissipation and the hyperbolicity are equally important, we expect that only part of the perturbation diffuses.

The differential system (11.1), (11.2) is of mixed type because the first block (11.1) are conservation laws, while the second (11.2) are balance laws. Therefore it is evident, that in order that the dissipation presented in the extended equation can also have a role in the block of conservation laws it is necessary to have a "coupling effect". The coupling condition is the so-called K-condition (Shizuta–Kawashima[37]) such that for a generic system of balance law of hyperbolic type with a convex extension there exist global smooth solutions. More precisely, provided the initial data are sufficiently smooth in L_2 norm, there exists a global smooth solution for all time (see Hanouzet and Natalini (2003)[38] in the case of one-dimensional space and Wen-An Yong[39] for generic space dimension). Moreover Ruggeri and Serre[40] have proved – under the same hypothesis – that the constant states are stable.

In Ruggeri (2004)[41] it was verified that the systems (11.1), (11.2) and (11.3) satisfy all the conditions of the previous theorems and therefore it is possible to conclude that if the initial data of an Extended Thermodynamics are sufficiently small then there exists a smooth global solution that converges to a constant solution of the equilibrium relativistic Euler fluids.

11.2 *Extended thermodynamics of moments*

Unfortunately in the classical framework it was also observed[42] that to have good agreement between theory and experiment (in particular sound waves in the limit of high frequencies or scattering of light) we need an ET with more moments. In this case it is not possible to construct a pure phenomenological theory as in the case of 14 moments and the problem becomes how it is possible to close the moment equations when we truncate the infinite hierarchy of moments. More precisely consider a finite number of moment equations until the tensorial index n:

$$\nabla_\alpha T^{\alpha A} = \pi^A, \quad A = 0, 1, 2, \ldots, n \tag{11.8}$$

[37] Shizuta, Y. and Kawashima, S. (1985) *Hokkaido Math. J.*, **14**, 249–75.
[38] Hanouzet, B. and Natalini, R. (2003)*Arch. Rat. Mech. Anal.*, **169**, 89.
[39] Yong, W. A. (2004) *Arch. Rat. Mech. Anal.*, **172**, 247.
[40] Ruggeri, T. and Serre, D. (2003) *Q. Appl. Math.*, **62**(1), 163.
[41] Ruggeri, T. (2004) *Il Nuovo Cimento B*, **119**(7–9), 809–21 (*Lecture Notes of the International Conference in honour of Y. Choquet-Bruhat: Analysis, Manifolds and Geometric Structures in Physics*, eds. G. Ferrarese and T. Ruggeri).
[42] Müller, I. and Ruggeri, T. (1998) *Rational Extended Thermodynamics*, Springer Tracts in Natural Philosophy, 37 (2nd Edition), Springer-Verlag, New York.

for the moments $T^{\alpha A}$ and productions π^A given by

$$T^{\alpha A}(x) = \int p^\alpha p^A f\, \omega_p, \quad \pi^A(x) = \int Q p^A f\, \omega_p, \tag{11.9}$$

where A is a multindex

$$p^A = \begin{cases} 1 \\ p^{\alpha_1} p^{\alpha_2} \dots p^{\alpha_A} \end{cases}, \quad F^{\alpha A} = \begin{cases} F^\alpha \\ F^{\alpha\alpha_1 \dots \alpha_A} \end{cases}, \quad \pi^A = \begin{cases} 0 & \text{for } A = 0 \\ \pi^{\alpha_1 \dots \alpha_A} & \text{for } A \geqslant 1 \end{cases}$$

and $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_A \leq 3$. We recall that the first five equations of (11.8) are the conservation laws of mass, momentum and energy; accordingly the first five productions vanish: $\pi^A = 0$ for $A = 0, 1$.

The closure procedure of the ET[43] assumes that any solution of the finite moment equations (11.8) with $A = 0, \dots, n$ is also a solution of the supplementary entropy law (11.5) with H^α given by the formula (10.1). We observe that this differs from the case of the infinite number of equations in which (11.5) is verified by the H-theorem, since in the present case this entropy inequality becomes an assumption and a constraint for the compatible distribution function. It is possible to prove that the "truncated" distribution function has the following expression

$$f_n = e^{\chi/k} \quad \text{where} \quad \chi = \sum_{A=0}^{n} u'_A(x) p^A$$

and the *main field* components u'_A are solutions of the symmetric hyperbolic system:

$$\sum_{B=0}^{n} H^{\alpha AB}(u'_C)\partial_\alpha u'_B = \pi^A(u'_C), \quad A = 0, 1 \dots n \tag{11.10}$$

where the symmetric matrices

$$H^{\alpha AB} = \int \frac{1}{k} e^{(\chi/k)} p^\alpha p^A p^B \omega_p \tag{11.11}$$

are such that $H^{\alpha AB}\xi_\alpha$ is a positive definite matrix for any timelike vector. This implies that the system of truncated moments (11.10) is symmetric hyperbolic.

Of course the f_n is not a solution of the Boltzmann equation and there is the conjecture (not yet proved) that for large n, f_n converges to f. An interesting indication of the validity of this conjecture is the asymptotic behaviour of the maximum characteristic velocity for large n.

11.3 *Maximum characteristic velocity*

The wave surface $\phi(x) = 0$ is the solution of the characteristic equation

$$\det(H^{\alpha AB}\partial_\alpha \phi) = 0.$$

[43] Boillat, G. and Ruggeri, T. (1999) *Cont. Mech. Thermodyn.*, **11**, 107–11.

As a consequence, the 4-gradient $\partial_\alpha \phi$ cannot be timelike and therefore the velocities of wave propagation cannot exceed the velocity of light. When the number of moments increases it has already been shown that the maximum wave velocity cannot decrease.[44] The question is now: does this velocity tends to c when n tends to infinity?

The wave velocity λ in the direction of the normal $n \equiv (n_i)$ to the wave front is an eigenvalue of

$$\det(H^{iAB} n_i - \lambda H^{0AB}) = 0 \tag{11.12}$$

Using the fact that the matrix in (11.12) is negative semi-definite for the maximum eigenvalue it is possible to prove the following inequality for a non-degenerate gas:

$$\frac{2n-1}{\gamma} \frac{K_{n+1}(\gamma)}{K_{n+2}(\gamma)} \leq \frac{\lambda_{\max}^2}{c^2} \leq 1, \tag{11.13}$$

where $\lambda_{\max}$ is the maximum characteristic velocity evaluated in an equilibrium state, and $\gamma = mc^2/kT$ and $K_n(\gamma)$ are the Bessel functions. Using the recurrence relation for Bessel functions $K_{n+2} - K_n = 2(n+1)K_{n+1}/\gamma$ it is possible to verify[45] that the limit value for $n \to \infty$ of the left-hand side of (11.13) is 1 and therefore the limit value of $\lambda_{\max}$ is c, since it has already been proved that it cannot be larger. Thus it is possible to conclude that when the number of moments tends to infinity the maximum velocity in equilibrium tends to the velocity of light. This result can be proved also for degenerate gases.[46] We observe that the proof is completely independent of the collisional operator Q in the Boltzmann equation. A numerical behaviour of the maximum velocity for increasing number of moments and also for a degenerate gas can be seen in Brini and Ruggeri (1999).[47]

[44] Boillat, G. and Ruggeri, T. (1997) *Arch. Rat. Mech. Anal.* **137**, 305–20.
[45] Boillat, G. and Ruggeri, T. (1999) *Cont. Mech. Thermodyn.*, **11**, 107–11.
[46] Boillat, G. and Ruggeri, T. (1999) *J. Math. Phys.*, **40**(12), 6399–404.
[47] Brini, F. and Ruggeri, T. (1999) *Cont. Mech. Thermodyn.*, **11**, 331–8.

XI

PROGRESSIVE WAVES

1 Introduction

We take afresh in this chapter the treatment of progressive[1] waves for non-linear equations used in Sections III.12 and III.13 to construct weak gravitational and electromagnetic waves on a given electrovac Einsteinian spacetime.

The fundamental improvement initiated by Leray in the construction of high-frequency waves for linear systems permits the extension to quasilinear systems[2], and the appearance of new properties linked to the non-linearities, in some sense similar to shocks[3]. There is in particular a distortion of signals, for instance for high-frequency waves in compressible fluids. We study general first-order quasilinear systems in Section III.3 and application to relativistic fluids in Section 4.

Many equations of mathematical physics are second-order quasilinear wave equations. We study this case in Section 5, showing in particular the role played by the null condition introduced independently by Christodoulou and Klainerman (see Chapter 15) in their study of the global existence of solutions of the Cauchy problem.

The fundamental field equations (standard model, Einstein equations) are quasilinear second-order partial differential equations, but are not well posed due to gauge invariance. In Section 6 we introduce a polarized "null condition". We show that it is satisfied by the standard model, but not quite by the Einstein equations. We construct for both these systems asymptotic high-frequency (progressive waves) solutions with linear transport law along the rays, a remarkable property in the case of the Einstein equations, which are quasilinear and not semilinear. However, in the case of the Einstein equations the wave inflicts a "back reaction" on the background metric. The linear propagation and the existence of the back reaction come from the fact that the Einstein equations satisfy almost, but not quite, the polarized null condition[4].

[1] Also called high-frequency waves, though for non-linear propagation the waves may be non-periodic.

[2] Choquet-Bruhat, Y. (1964) *C. R. Acad. Sci. Paris*, **258**, 3809–12 and (1969) *J. Math. Pures Appl.*, **48**, 117–58.

[3] See Boillat, G. (1996) in *Recent Mathematical Methods in Non-Linear Wave Propagation*, Lecture Notes in Mathematics 1640, Springer, Berlin.

[4] See further comments in Section XV.7.

2 Quasilinear systems

The WKB approximation, or its generalization by Lax to asymptotic series, cannot take care of non-linear effects because their non-linear combinations are not of the same type.

Leray's definition formally applies to non-linear equations, but it is not possible in general to obtain asymptotic series with coefficients bounded at all orders due to the appearance of products of derivatives with respect to ξ which do not satisfy the boundedness *ansatz* (proposition 3 of Section III.12), which we recall

Proposition 2.1 *A C^1 function $\xi \mapsto f(\xi)$ can be uniformly bounded in $\xi \in R$ only if its derivative*

$$f' := \frac{\partial f}{\partial \xi} \tag{2.1}$$

has a uniformly bounded integral on any interval (ξ_0, ξ). In particular f' cannot be independent of ξ.

However, for quasilinear systems, it is possible to construct truncated series which define approximate solutions in a well-defined sense by expanding in Taylor series the non-linear terms around given values.

We consider fields on a manifold V and give the following definitions.

Definition 2.2 *A field $u(x, \omega\phi(x))$ is called a **high-frequency wave**, or a **progressive wave**, of order p if there exists a background field $\underline{u}(x)$ and fields $u_p(x, \omega\phi(x))$ and $u_{p+1}(x, \omega\phi(x))$ such that*

$$u(x, \omega\phi(x)) \equiv \underline{u}(x) + \omega^{-p}\{u_p(x, \xi) + \omega^{-1}u_{p+1}(x, \xi)\}_{\xi=\omega\phi(x)}, \tag{2.2}$$

where u_p and u_{p+1} are bounded in some norm for $x \in V$, uniformly in $\xi \in R$.

Let $N(u)$ be a non-linear differential operator of order m, depending smoothly on u and its derivatives up to order m in some domain of values of these quantities containing the background, i.e. $\underline{u}$ and its derivatives up to order m. One computes $N(u(x, \omega\phi(x))$ using expansion in the unknown around the background and the derivation formula (III.12.4).

Definition 2.3 *A progressive wave u is called an asymptotic solution of order q of a non-linear operator $N(u) = 0$ on the manifold V if*

$$N(u(x, \omega\phi(x)) \equiv \omega^{-q}F(x, \omega) \tag{2.3}$$

where $F(x, \omega)$ is bounded in norm for $x \in V$, uniformly in $\omega > 0$.

An asymptotic solution gives an approximate solution for a large value of ω.

3 Quasilinear first-order systems

First-order systems are simpler to treat in general than arbitrary systems. We give the outline of this treatment. We will apply it in the following section to the construction of progressive waves for relativistic fluids, possibly coupled to weak progressive gravitational waves.

3.1 *Phase and polarization*

We consider a first-order system of quasilinear differential equations for an unknown u,

$$a(x,u)\partial u + f(x,u) = 0. \tag{3.1}$$

We work in local coordinates x^α; u is represented by a set of N scalar functions u^I and a by matrices $a^{I,\alpha}_J$. The system reads

$$a^{I,\alpha}_J(x,u)\partial_\alpha u^J + f^I(x,u) = 0. \tag{3.2}$$

We expand the coefficients in Taylor series around a background $\underline{u}$, i.e.

$$a^{I,\alpha}_J(u) = \underline{a}^{I,\alpha}_J + \underline{\partial a^{I,\alpha}_J/\partial u^H}(u^H - \underline{u}^H) + \cdots \tag{3.3}$$

and a similar expansion for f, where we underline the value taken by a function of u for $u = \underline{u}$.

We look for a progressive wave $u - \underline{u}$ of order 1

$$u(x,\xi) = \underline{u}(x) + \omega^{-1}u^{(1)}(x,\xi) + \omega^{-2}u^{(2)}(x,\xi), \quad \xi = \omega\phi(x), \tag{3.4}$$

asymptotic solution of (3.1) in the sense that

$$a(x,u)\partial u + f(x,u) = O(\omega^{-2}). \tag{3.5}$$

We use for a function of x and ξ the notation

$$\underline{\partial}_\alpha u(x,\omega\phi(x) := \left\{\frac{\partial}{\partial x^\alpha}u(x,\xi)\right\}_{\xi=\omega\phi(x)}, \quad u'(x,\omega\phi(x) := \left\{\frac{\partial}{\partial\xi}u(x,\xi)\right\}_{\xi=\omega\phi(x)}. \tag{3.6}$$

We have the identity

$$\partial_\alpha u \equiv \underline{\partial}_\alpha u + \omega\phi_\alpha u', \phi_\alpha := \frac{\partial\phi}{\partial x^\alpha} \tag{3.7}$$

Reporting the expansion (3.4) in Equation (3.1) and using the Taylor formula (3.3) the system (3.1) becomes

$$\underline{a^{I,\alpha}_J}(\partial_\alpha\underline{u}^J + \phi_\alpha u'^{(1)J}) + \underline{f}^I + \omega^{-1}(\underline{a}^{I,\alpha}_J\phi_\alpha u'^{(2)J} + F^I) + \omega^{-2}\mathcal{R} = 0, \tag{3.8}$$

with

$$F^I := \underline{a}^{I,\alpha}_J\underline{\partial}_\alpha u^{(1)J} + \underline{\partial a^{I,\alpha}_J/\partial u^H}u^{(1)H}(\partial_\alpha\underline{u}^J + \phi_\alpha u'^{(1)J}) + \underline{\partial f^I/\partial u^H}u^{(1)H}. \tag{3.9}$$

The coefficient $\mathcal{R}$ of the remainder is bounded under smoothness conditions on the coefficients of the original system (3.1).

The progressive wave u will be an asymptotic solution if the coefficients of ω^0 and ω^{-1} in (3.8) vanish. The condition of order zero is

$$\underline{a_J^{I,\alpha}}(\partial_\alpha \underline{u}^J + \phi_\alpha u'^{(1)J}) + \underline{f}^I = 0 \tag{3.10}$$

This equation and the averaging *ansatz*, Proposition 2.1, imply that the background $\underline{u}$ satisfies the original system, since by (3.10) this original operator is equal to a derivative with respect to ξ. The system (3.10) reduces then to the linear homogeneous system for $u^{(1)\prime}$

$$\underline{a_J^{I,\alpha}}\phi_\alpha u'^{(1)J} = 0. \tag{3.11}$$

This system has a non-zero solution $u^{(1)\prime}$ if and only if ϕ_α is a solution of the **characteristic equation** of the background

$$\underline{D}(\phi) := Det(\underline{a_J^{I,\alpha}}\phi_\alpha) = 0. \tag{3.12}$$

In other words, zero must be an eigenvalue of the characteristic matrix $(\underline{a_J^{I,\alpha}}\phi_\alpha)$. The form of the solution $u^{(1)\prime}$ depends on the multiplicity of this eigenvalue and the number of eigenvectors. If there are k linearly independent eigenvectors $X_{(p)}$ corresponding to the phase ϕ then the general solution of (3.11) is

$$u^{(1)\prime}(x,\xi) = \sum_{p=1,\dots k} V_{(p)}(x,\xi) X_{(p)}(x) \tag{3.13}$$

where the $V_{(p)}$ are arbitrary functions of x and ξ. We call this formula the **polarization condition** for the wave.

3.2 *Propagation equations*

The condition of order 1, corresponding to the vanishing of the coefficient of ω^{-1} in (3.8), is

$$\underline{a}_J^{I,\alpha}\phi_\alpha u^{'(2)J} + F^I = 0 \tag{3.14}$$

Elementary algebra shows that necessary and sufficient[5] conditions for the existence of a solution $u^{'(2)}$ of the above system are

$$Y_I F^I = 0, \tag{3.15}$$

where Y is any of the left eigenvectors with eigenvalue zero of the background characteristic matrix.

[5] The condition is necessary already for a zero $u^{(2)}$. It is also sufficient if one admits a non-zero $u^{(2)}$ in the asymptotic expansion (3.4); this fact is the motivation for the introduction of this term in the expansion.

3.2.1 *Simple characteristic*

Suppose the rank of the $N \times N$ background characteristic matrix is $N - 1$; that is, $grad\phi := (\phi_\alpha)$ is a simple root of the characteristic equation. Then there is only one, up to a scale factor, right (respectively left) eigenvector $X(x)$ corresponding to the eigenvalue zero. The general value of $u'^{(1)I}$ is, after choosing the eigenvector X, determined up to a scalar factor $U(x, \xi)$. Choosing 0 as an additional term in integration with respect to ξ, the polarized wave is

$$u^{(1)I}(x, \xi) = U(x, \xi) X^I(x). \tag{3.16}$$

We report this value of $u^{(1)}$ in Equation (3.15), after choosing the left eigenvector Y. We obtain for U a semilinear partial differential equation in x and ξ which reads after reorganization of the terms

$$Y_I \underline{a}_J^{I,\alpha} X^J \frac{\partial U}{\partial x^\alpha} + \Phi U \frac{\partial U}{\partial \xi} + \Psi U = 0 \tag{3.17}$$

where Φ and Ψ are functions of x given by

$$\Phi := Y_I \frac{\partial a_J^{I,\alpha}}{\underline{\partial u^H}} \phi_\alpha X^H X^J, \tag{3.18}$$

$$\Psi := Y_I \left\{ \underline{a}_J^{I,\alpha} \frac{\partial X^J}{\partial x^\alpha} + \frac{\partial a_J^{I,\alpha}}{\underline{\partial u^H}} X^H \partial_\alpha \underline{u}^J + \frac{\partial f^I}{\underline{\partial u^H}} X^H \right\}. \tag{3.19}$$

The coefficient Ψ vanishes if the background is constant and if f does not depend on u.

The following lemmas give simple expressions for the coefficient of the linear derivative of U and for Φ.

Lemma 3.1 *There is a scalar k depending on the choice of X and Y such that*

$$Y_I \underline{a}_J^{I,\alpha} X^J \equiv k \underline{A}^\alpha \tag{3.20}$$

where

$$\underline{A}^\alpha := \frac{\partial \underline{D}(\phi)}{\partial \phi_\alpha} \tag{3.21}$$

$\underline{D}(\phi)$ being the characteristic determinant with elements $\underline{a}_J^{I,\alpha} \phi_\alpha$; that is, $\underline{A}^\alpha$ is tangent to the ray associated with the characteristic ϕ.

Lemma 3.2 *It holds that*

$$\Phi \equiv k \frac{\partial \underline{D}(\phi)}{\partial \underline{u}^H} X^H \tag{3.22}$$

Proof

1. By elementary algebra, it is possible to choose the vectors X and Y such that

$$X^I Y_J = k M_J^I \tag{3.23}$$

where M_J^I is the cofactor of the element $\underline{D}_I^J \equiv \underline{a}_J^{I,\alpha}\phi_\alpha$ in $\underline{D}(\phi)$. The law of derivation of a determinant gives

$$\underline{A}^\alpha := \frac{\partial \underline{D}(\phi)}{\partial \phi_\alpha} = M_I^J \frac{\partial \underline{a}_J^{I,\alpha}\phi_\alpha}{\partial \phi_\alpha} = M_I^J \underline{a}_J^{I,\alpha}. \tag{3.24}$$

2. The same argument gives

$$\Phi := Y_I \frac{\partial a_J^{I,\alpha}}{\partial \underline{u}^H}\phi_\alpha X^H X^J = k M_I^J \frac{\partial a_J^{I,\alpha}\phi_\alpha}{\partial \underline{u}^H} X^H \equiv k \frac{\partial \underline{D}(\phi)}{\partial \underline{u}^H} X^H. \tag{3.25}$$

□

Characteristic manifolds such that $\Phi \equiv 0$ have been called by Lax and Boillat in their study of shock waves **exceptional characteristic** manifolds.

The integration of the first-order quasilinear partial differential equation (3.17) is done, classically[6], through the integration of its bicharacteristic system, a system of ordinary differential equations which are here

$$d\lambda = \frac{dx^\alpha}{k\underline{A}^\alpha} = \frac{d\xi}{\Phi U} = -\frac{dU}{\Psi}. \tag{3.26}$$

The solution of the differential equations

$$d\lambda = \frac{dx^\alpha}{k\underline{A}^\alpha} \tag{3.27}$$

such that the coordinates x^α of x take given values y^α for $\lambda = 0$ is the ray in spacetime V passing through the point y,

$$\lambda \mapsto x(\lambda, y) \quad \text{with} \quad x(0, y) = y. \tag{3.28}$$

Let $s(y) = 0$ be a submanifold Σ of spacetime. Let $W(y, \eta)$ be a given function on $V \times R$. The solution $U(x, \xi)$ of the partial differential equation (3.17) such that

$$U(x, \xi)|_{x=y, \xi=\eta} = W(y, \eta) \quad \text{when } s(y) = 0 \tag{3.29}$$

is generated by the rays issuing from Σ; that is, obtained by elimination of λ, y, and η between the solutions of the differential system (3.26)

$$x = x(\lambda, y), \quad \xi = \xi(\lambda, y, \eta), \quad U = U(\lambda, y, \eta) \tag{3.30}$$

which satisfy

$$x(0, y) = y, \quad \xi(0, y, \eta) = \eta, \quad U(0, y, \eta) = W(y, \eta). \tag{3.31}$$

Under differentiability assumptions and if the submanifold Σ is transversal to the rays, equations (3.30) can be inverted for small enough λ. However, independently of singularities due to the possible formation of caustics of the rays, the

[6] See e.g. CB-DM1 IV

presence of U, a function of x but also of ξ, in the quotient $d\xi/\Phi U$ prevents the solution U from being a product

$$U(x,\xi) = U_1(x)U_2(\xi) \tag{3.32}$$

for initial values of the form

$$W(y,\eta) = W_1(y)W_2(\eta); \tag{3.33}$$

that is, for non-exceptional waves there is a **distortion of signals**[7]. It may lead to singularities.

3.2.2 *Multiple characteristic*

The properties of progressive waves with a phase ϕ whose gradient is a multiple root of the characteristic determinant depends on the corresponding eigenvectors.

If the number k of linearly independent eigenvectors is equal to the multiplicity of the root, a property which is always true when the system is symmetrizable hyperbolic[8], then it can be proved that the system admits a progressive wave, asymptotic solution perturbation of the considered background, which is of the form

$$u^{(1)}(x,\xi) = \sum_{p=1,\ldots,k} U_{(p)}(x,\xi)h_{(p)}(x). \tag{3.34}$$

The functions $U_{(p)}$ obey a system of ordinary differential first-order equations: such multiple characteristics are always exceptional[9].

When *the number k of independent eigenvectors is less than the multiplicity of the root of the characteristic determinant*, there can still be asymptotic waves of the form (3.34), but under restrictive conditions. Consider the case where $grad\phi$ is a root of multiplicity 2 of the characteristic polynomial to which there corresponds only one eigenvector X. It is proved that there exists a polynomial called **subcharacteristic**[10], depending on the background values of the principal and not principal coefficients of the operator, which must also admit $grad\phi$ as a root for an asymptotic progressive wave solution to exist. This property is also sufficient for the existence of a function $U(x,\xi)$ such that

$$U(x,\xi)X(x) \tag{3.35}$$

is an asymptotic wave of order zero for the system, at the chosen background. It is an asymptotic wave of order one if the function U satisfies a second-order partial differential equation in x and ξ[11].

[7] See Choquet-Bruhat, Y. (1969) *J. Math. Pures Appl.*, **48**, 117–58, part V, referred to below as CB69.

[8] See Remark 4.5 in Appendix IV.

[9] General result due to G. Boillat.

[10] Introduced by J. Leray in linear systems, and studied and used in the general framework of the Cauchy problem by Vaillant, J. (1968) *J. Math. Pures Appl.*, **47**, 1–40.

[11] See more the elaborate treatment in CB69, sections 5 and 6.

4 Progressive waves in relativistic fluids

The equations of perfect relativistic fluids (see Chapter 9) are a well-posed system of first-order quasilinear partial differential equations to which the general method of construction of progressive waves have been applied in a given spacetime in CB69. We prove in this section some general results about progressive waves in relativistic fluids and their coupling with weak gravitational waves. Results about charged fluids can be found in CB69. The Leray–Ohya non-strict hyperbolicity of fluids with finite electric conductivity is illustrated by the appearance of second-order differential equations in high-frequency wave propagation. High-frequency waves in fluids with infinite conductivity have been studied by Anile and Greco[12]. The case of the coupling of fluid waves with strong gravitational waves is treated by Choquet-Bruhat and Greco in an article reproduced in this volume.

4.1 *Equations*

Recall that the Euler equations (IX.4.2, IX.4.3) satisfied by a perfect fluid in a spacetime (V, g) are

$$M_c := (\mu + p)\nabla_\alpha u^\alpha + u^\alpha \nabla_\alpha \mu = 0 \quad \text{continuity equation,} \tag{4.1}$$

$$M^\beta := (\mu + p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\alpha u^\beta)\nabla_\alpha p = 0, \quad \text{equations of motion.} \tag{4.2}$$

For simplicity of writing we consider isentropic motion and a barotropic equation of state[13]

$$p = p(\mu). \tag{4.3}$$

The Einstein equations with this fluid as a source are, in a four-dimensional spacetime[14],

$$R_{\alpha\beta} = \rho_{\alpha\beta}, \quad \rho_{\alpha\beta} \equiv (\mu + p)u_\alpha u_\beta + \frac{1}{2}(\mu - p)g_{\alpha\beta}. \tag{4.4}$$

4.2 *Progressive waves*

We look for an asymptotic solution of the coupled Einstein–Euler equations as a coupled perturbation of a background $(\underline{u}, \underline{\mu}, \underline{g})$ by a progressive fluid wave and a weak gravitational wave.

[12] See e.g. the book by Anile, M. (1987) *Relativistic Hydrodynamics and Magnetohydrodynamics*, Cambridge University Press.

[13] The general case is treated in CB69.

[14] Analogous results hold in any dimension. We choose $n = 3$ for simplicity of writing, and because it is the most interesting case for physical applications.

The Lorentzian metric, weak[15] gravitational wave perturbation of a background $\underline{g}$, is

$$g_{\alpha\beta}(x) = \underline{g}_{\alpha\beta}(x) + \{\omega^{-2} v_{\alpha\beta}(x,\xi) + \omega^{-3} w_{\alpha\beta}(x,\xi)\}_{\xi=\omega\phi(x)}. \tag{4.5}$$

The fluid velocity and energy density are of the form

$$u^{\alpha}(x) = \underline{u}^{\alpha}(x) + \{\omega^{-1} u^{(1)\alpha}(x,\xi) + \omega^{-2} u^{(2)\alpha}(x,\xi)\}_{\xi=\omega\phi(x)}, \tag{4.6}$$

$$\mu(x) = \underline{\mu}(x) + \{\omega^{-1}\mu^{(1)}(x,\xi) + \omega^{-2}\mu^{(2)}(x,\xi)\}_{\xi=\omega\phi(x)}. \tag{4.7}$$

with all quantities uniformly bounded in ξ. These expansions imply that

$$p(x) = \underline{p}(x) + \omega^{-1} p^{(1)}(x,\xi) + \omega^{-2} p^{(2)}(x,\xi). \tag{4.8}$$

with, using the Taylor formula

$$p^{(1)}(x,\xi) := \underline{p'_{\mu}}(x)\mu^{(1)}(x,\xi), \quad \text{with} \quad p'_{\mu} := \frac{dp}{d\mu.} \ . \tag{4.9}$$

The source tensor ρ admits therefore the expansion

$$\rho_{\alpha\beta} = \underline{\rho}_{\alpha\beta} + \omega^{-1}\left\{\underline{(\mu+p)}(\underline{u}_{\beta}u^{(1)}{}_{\alpha} + \underline{u}_{\alpha}u^{(1)}{}_{\beta}) + \underline{u}_{\alpha}\underline{u}_{\beta}(p^{(1)} + \mu^{(1)}) + \frac{1}{2}\underline{g}_{\alpha\beta}(\mu - p)^{(1)}\right\} + O(\omega^{-2}). \tag{4.10}$$

4.3 *Phases and polarizations*

The Christoffel symbols of the weak gravitational wave are (Section III.12)

$$\Gamma^{\lambda}_{\alpha\beta} = \underline{\Gamma}^{\lambda}_{\alpha\beta} + \omega^{-1}\gamma^{\lambda}_{\alpha\beta} + O(\omega^{-2}), \tag{4.11}$$

The principal terms of the Einstein equations with source ρ are in ω^0 and read (see Section III.12)

$$\underline{R}_{\alpha\beta} - \frac{1}{2}\{\underline{g}^{\lambda\mu}\phi_{\lambda}\phi_{\mu}v''_{\alpha\beta} - \phi_{\beta}P''_{\alpha} - \phi_{\alpha}P''_{\beta}\}(x,\omega\phi(x)) = \underline{\rho}_{\alpha\beta}. \tag{4.12}$$

with

$$P''_{\alpha} := \phi^{\mu}\left(v''_{\mu\alpha} - \frac{1}{2}\underline{g}_{\mu\alpha}v''^{\rho}_{\rho}\right), \tag{4.13}$$

The averaging *ansatz* implies that

$$\underline{R}_{\alpha\beta} - \underline{\rho}_{\alpha\beta} = 0. \tag{4.14}$$

The equations (4.12) reduce to

$$\underline{g}^{\lambda\mu}\phi_{\lambda}\phi_{\mu}v''_{\alpha\beta} - \phi_{\beta}P''_{\alpha} - \phi_{\alpha}P''_{\beta} = 0. \tag{4.15}$$

They do not contain the fluid variables.

[15] See Section III.12.

The same conclusions for ϕ and v hold as in Section IV.12, namely ϕ isotropic for $\underline{g}$ and v polarized as in (III.12.18) or ϕ non-isotropic and a v which we called non-significant because it can be made to vanish by a change of reference frame. However in the presence of matter there are privileged frames attached to it. In these frames "gauge" gravitational progressive waves with non-isotropic phase can be observed, and we will take them into account.

When $grad\phi$ is not isotropic Equations (4.15) are equivalent to

$$v_{\alpha\beta} = \phi_\beta P_\alpha + \phi_\alpha P_\beta \tag{4.16}$$

with p an arbitrary vector function of x and ξ which we call the **polarization vector**.

The principal terms of the fluid equations are also in ω^0, sum of terms independent of ξ and terms linear in the derivatives with respect to ξ of the perturbation. The averaging *ansatz* implies that the first set of terms vanishes; that is, the background fluid must satisfy the fluid equations in the background metric. The annulation of the second set is a 5×5 linear homogeneous system for $u^{(1)\alpha\prime}$ and $p^{(1)\prime}$ for the set of unknowns. We use the notation $(u'^I) := (u^{(1)\alpha\prime}, p^{(1)\prime})$ for the unknowns and $M^J_I u'^I = 0$, $I, J = 0, 1, 2, 3, 4$ for the homogeneous system they satisfy. These equations are

$$M^4_I u'^I \equiv \underline{(\mu + p)}\phi_\alpha u^{(1)\prime\alpha} + \underline{u^\alpha}\phi_\alpha \mu^{(1)\prime} = 0, \tag{4.17}$$

$$M^\beta_I u'^I \equiv \underline{(\mu + p)u^\alpha}\phi_\alpha u^{(1)\prime\beta} + \underline{(g^{\alpha\beta} + u^\alpha u^\beta)}\phi_\alpha \underline{\frac{\partial p}{\partial \mu}} \mu^{(1)\prime} = 0, \tag{4.18}$$

They do not contain the gravitational perturbation.

The determinant D of the matrix M coincides with the determinant computed in Section IX.8 when looking for the wave fronts of a perfect fluid; that is, in the notation of the present chapter, we have

$$D \equiv -(\underline{\mu + p})^3 (\underline{u}^\alpha \phi_\alpha)^3 D_A, \tag{4.19}$$

with

$$D_A := (1 - \underline{p'_\mu})(u^\alpha \phi_\alpha)^2 - \underline{p'_\mu}\phi^\alpha \phi_\alpha, \qquad \underline{p'_\mu} := \underline{\frac{dp}{d\mu}}. \tag{4.20}$$

The determinant D vanishes, and the system (4.17), (4.18) admits non-zero solutions if and only if $\phi =$ constant is a fluid wave front. There are two types of wave front:

$$D_A = 0, \quad \text{acoustic wave fronts} \tag{4.21}$$

and

$$\underline{u}^\alpha \phi_\alpha = 0 \quad \text{matter wave fronts.} \tag{4.22}$$

We see that the fluid wave is zero at first order if the phase is isotropic, except if the fluid is incompressible, i.e. except if the equation of state is $p = \mu$.

4.4 *Polarization and propagation of acoustic waves*

When the fluid is compressible[16] the acoustic wave fronts are non-isotropic. The gravitational wave is then of the form (4.16).

4.4.1 *Polarization*

If $grad\phi$ is normal to an acoustic wave front, the 5×5 determinant D vanishes and is of rank 4. The general solution $(u'^A) := (u^{(1)\alpha\prime}, u^{(1)4\prime} := p^{(1)\prime})$ of Equations (3.11) (the polarization conditions) is therefore determined up to a proportionality factor. Taking the primitive with respect to ξ and neglecting the addition of terms independent of ξ, we have

$$u^{(1)I}(x,\xi) = U(x,\xi)X^I(x) \tag{4.23}$$

where (X^I) is a particular solution of the system (3.11) for the chosen ϕ (recall that the coefficients of the system do not depend on ξ). Such a solution is proportional to the cofactors of the elements of a given line of D. A possible choice is (index raised with $\underline{g}$)

$$X^\alpha = \underline{p'_\mu}\ \phi^\alpha, \quad X^4 = \underline{\mu + p}\underline{u}^\alpha\phi_\alpha. \tag{4.24}$$

4.4.2 *Propagation*

The propagation of the scale factor U is deduced from the asymptotic equations of order 1; that is, for the continuity equation

$$\begin{aligned} M^{(1)4} := {} & \underline{\mu + p}\underline{\nabla}_\alpha u^{(1)\alpha} + (\underline{\nabla}_\alpha \underline{u^\alpha} + \phi_\alpha u^{(1)\alpha\prime})(\mu^{(1)} + p^{(1)}) + (\underline{u}^\alpha + u^{(1)\alpha\prime})\partial_\alpha\mu^{(1)} \\ & + u^{(1)\alpha}(\partial_\alpha\underline{\mu} + \phi_\alpha\mu^{(1)\prime}) + (\underline{\mu + p})\underline{u}^\beta\Gamma^{\alpha(1)}_{\alpha\beta} = f^4, \end{aligned} \tag{4.25}$$

with

$$f^4 := -\{\underline{(\mu + p)}\phi_\alpha u^{(2)\prime\alpha} + \underline{u^\alpha}\phi_\alpha\mu^{(2)\prime}\} \tag{4.26}$$

and for the equations of motion

$$\begin{aligned} M^{\beta(1)} := {} & \underline{\mu + p}\underline{u}^\alpha\underline{\nabla}_\alpha u^{(1)\beta} + (\mu^{(1)} + p^{(1)})\underline{u}^\alpha(\underline{\nabla}_\alpha\underline{u}^\beta + \phi_\alpha u^{(1)\beta\prime}) \\ & + (\partial_\alpha\underline{p} + p^{(1)\prime}\phi_\alpha)(\underline{u}^\alpha u^{(1)\beta} + \underline{u}^\beta u^{(1)\alpha}) + \underline{\mu + p}\underline{u}^\alpha\underline{u}^\gamma\Gamma^{(1)\beta}_{\alpha\gamma} = f^\beta \end{aligned} \tag{4.27}$$

with

$$f^\beta := -\{\underline{(\mu + p)u^\alpha}\phi_\alpha u^{(2)\prime\beta} + \underline{(g^{\alpha\beta} + u^\alpha u^\beta)}\phi_\alpha p^{(2)\prime}\} \tag{4.28}$$

The right-hand sides of these equations are the action of the matrix M on $(u^{(2)I\prime})$,

$$f^I \equiv M^I_J u^{(2)I\prime}. \tag{4.29}$$

[16] Incompressible fluids are included in the reproduced CB-Greco article.

Therefore, there exists $(u^{(2)I\prime})$ satisfying these equations if and only if[17] the left-hand side is orthogonal to the left eigenvector of the matrix M, i.e.

$$Y_I M^{(1)I} = 0, \quad \text{with} \quad Y_I M^I_J = 0. \tag{4.30}$$

The vector Y_I is found to be, up to a common factor

$$Y_4 = -\underline{u}^\alpha \phi_\alpha, \qquad Y_\alpha = \phi_\alpha. \tag{4.31}$$

We replace in Equations (4.30) the unknowns $u^{(1)I}$ by their polarized values

$$u^{(1)I} = UX^I, \quad \text{hence} \quad u^{(1)I\prime} \equiv \frac{\partial U}{\partial \xi} X^I. \tag{4.32}$$

We thus obtain for the scalar function U a non-linear first-order partial differential equation in x and ξ which is found to be

$$-k \left\{ \bar{A}^\alpha \frac{\partial}{\partial x^\alpha} U + \Phi U \frac{\partial U}{\partial \xi} + \Psi U \right\} = \Xi. \tag{4.33}$$

The left-hand side is the same as that found in CB69 with a given background metric; the functions Φ and Ψ depend on the backgrounds, fluid, and metric, and on the choice of the phase ϕ; we set to simplify notations,

$$a := \underline{u}^\alpha \phi_\alpha, \qquad \underline{p''_{\mu^2}} := \underline{\frac{d^2 p}{(d\mu)^2}}. \tag{4.34}$$

Then the coefficients in (4.33) are functions independent of ξ which read

$$k := -2a^2 \underline{\mu + p}, \quad \text{with} \quad a := \underline{u}^\alpha \phi_\alpha, \tag{4.35}$$

$$A^\alpha := (\, \underline{p'_\mu} - 1) \underline{u}^\alpha (\underline{u}^\beta \phi_\beta) + \underline{p'_\mu} \phi^\alpha \equiv \frac{1}{2} \frac{\partial D_A}{\partial \phi_\alpha}, \quad A^\alpha \text{ is the acoustic ray,} \tag{4.36}$$

$$\Phi := -a^2 \left\{ (1 - \underline{p'_\mu}) + \frac{p + \mu}{2\underline{p'_\mu}} \underline{p''_{\mu^2}} \right\} \tag{4.37}$$

and

$$\Psi := \frac{1}{2} (\Box_A \phi + a \underline{\nabla_\alpha u^\alpha} + \tilde{\Psi}), \tag{4.38}$$

where $\Box_A$ is the acoustic wave operator; that is,

$$\Box_A := \frac{\partial A^\alpha}{\partial \phi_\beta} \underline{\nabla_\alpha} \partial_\beta. \tag{4.39}$$

[17] The condition is obviously necessary. It is sufficient for the existence of some $u^{(2)I}$, not for $u^{(2)I} = 0$.

The additional function $\tilde{\Psi}$ vanishes if the background has a constant density[18], i.e. $\partial_\alpha \underline{\mu} = 0$.

The factor Φ determines the distortion of signals. It vanishes if the equation of state satisfies the differential equation

$$(\underline{p+\mu})\underline{p}''_{\mu} + 2\underline{p}'_{\mu}(1 - \underline{p}'_{\mu}). \tag{4.40}$$

Theorem 4.1 *Signals propagate without distortion in a relativistic fluid if it is incompressible or if it admits the equation of state*

$$p = -\frac{C}{\mu}, \quad C \text{ a constant.} \tag{4.41}$$

Proof For an incompressible fluid holds $\underline{p}'_{\mu} = 1$; (4.10).

Elementary computation shows that (4.41) also satisfies (4.40). Such waves are real ($p'_\mu > 0$) only for nm classical fluids ($p \, or \, \mu < 0$). □

The right-hand side Ξ of (4.33) is generated by the gravitational "gauge" wave,

$$v_{\alpha\beta} \equiv \phi_\alpha P_\beta + \phi_\beta P_\alpha; \tag{4.42}$$

it reads

$$\Xi \equiv -(\underline{\mu+p})\{Y_4 \underline{u}^\beta \Gamma^{\alpha(1)}_{\alpha\beta} + Y_\lambda \underline{u}^\alpha \underline{u}^\beta \Gamma^{(1)\lambda}_{\alpha\beta}\}. \tag{4.43}$$

We have (see Section III.12)

$$\Gamma^{(1)\lambda}_{\alpha\beta} = \frac{1}{2}(\phi_\alpha v'^{\lambda}_{\beta} + \phi_\beta v'^{\lambda}_{\alpha} - \phi^\lambda v'_{\alpha\beta}) = \phi_\alpha \phi_\beta P'^{\lambda} - \frac{1}{2}\phi^\lambda(\phi_\alpha P'_\beta + \phi_\beta P'_\alpha), \tag{4.44}$$

$$\Gamma^{(1)\alpha}_{\alpha\beta} = \frac{1}{2}\phi_\beta v'^{\lambda}_{\alpha} = \phi_\beta \phi_\alpha P'^{\alpha}. \tag{4.45}$$

We report these values in (4.43) and we find in the case of an acoustic wave, where $Y_4 = -\underline{u}^\alpha \phi_\alpha$, $Y_\alpha = \phi_\alpha$,

$$\Xi = (\underline{\mu+p})(\phi_\lambda \underline{u}^\lambda)\phi_\lambda \phi^\lambda \underline{u}^\beta P'_\beta. \tag{4.46}$$

We see that Ξ does not vanish if $\phi_\lambda \phi^\lambda \neq 0$ and $\underline{u}^\beta P'_\beta \neq 0$. We have proved the following theorem.

Theorem 4.2 *A weak gravitational wave with acoustic phase in a compressible fluid generates an acoustic wave if the gravitational polarization vector p is not orthogonal to the background fluid velocity.*

[18] See the full expression, with an equation of state depending also on entropy in CB69.

4.5 *Polarization and propagation of matter waves*

The matter wave phases are solutions of the linear equation $\underline{u}^{\alpha}\phi_{\alpha} = 0$ and are triple roots of the determinant D. For such a ϕ the determinant D is of rank 2. There are three linearly independent[19] vectors X which are solutions of the linear system (4.17), (4.18). One can choose, with $p = 1, 2, 3$

$$X^4_{(p)} = 0, \quad X^{\alpha}_{(p)} : \quad 3 \text{ linearly independent solutions of } X^{\alpha}_{(p)}\phi_{\alpha} = 0, \tag{4.47}$$

and for vectors Y

$$Y^{(p)}_4 = 0, \quad Y^{(p)}_{\alpha} := \underline{g}_{\alpha\beta} X^{\beta}_{(p)}. \tag{4.48}$$

The solution $u^{(1)}$ is of the form

$$u^{(1)I} = \sum_{p=1,2,3} U_{(p)}(x, \xi) X^I_{(p)}(x). \tag{4.49}$$

A first-order differential system of three equations for the three functions $U_{(p)}$ is obtained by reporting these values in the system given in Section 4.4.2. (Propagation) and making the three contracted products by the vectors $Y_{(q)}$, $q = 1, 2, 3$.

We find that no term $U_{(p)}\partial U_{(q)}/\partial\xi$ appears in the differential system[20], and the matter waves are exceptional in the sense of Lax–Boillat. This result is in agreement with the Boillat general theorem.

4.6 *"Gauge" gravitational waves*

We leave open for the reader the problem of writing propagation equations for "gauge" gravitational waves.

5 Quasilinear quasidiagonal second-order systems

The construction of asymptotic progressive waves by the GKL method applies to linear systems of arbitrary order[21]. It is possible to extend to general quasilinear systems the construction we gave in the previous sections of progressive waves which are asymptotic solutions of quasilinear first-order systems. We treat now quasilinear second-order systems. Such systems occur in many physical problems.

5.1 *Definitions*

A second-order system with unknown a set of tensor fields u on a C^{∞} manifold V reads

$$F(x, u, Du, D^2u) = 0,$$

[19] This property can be foreseen from the Leray hyperbolicity of the equations of relativist perfect fluids, but does not prove it.

[20] Choquet-Bruhat (1969).

[21] Such systems can always be written as first-order systems by introducing as new unknowns derivatives of the original ones. But it is a cumbersome procedure which may obscure physical properties.

where D is the covariant derivative with respect to a given metric on V. If the system is quasilinear, then F is linear in the second derivative of u; such a quasi-diagonal system reads in index notation

$$F^A(x,u,Du,D^2u) \equiv g^{\alpha\beta}(x,u,Du)D^2_{\alpha\beta}u^A + f^A(x,u,Du)) = 0, \tag{5.1}$$

with $u \equiv (u^A), A = 1,\dots,N$, and x^α local coordinates on V.

The Einstein–Maxwell (or Yang–Mills) scalar field equations are of this type when the metric is in wave gauge and the Yang–Mills potential is in Lorentz gauge.

We look for a progressive wave of order 1; that is, of the form[22]

$$u(x,\omega\phi(x)) := \underline{u}(x) + \{\omega^{-1}v(x,\xi)\}_{\xi=\omega\phi(x)}, \tag{5.2}$$

with $\underline{u}$ a tensor field on V, called background, and v a tensor field of the same type as $\underline{u}$, but depending on a real parameter $\xi \in R$. The tensor u given by (5.2) is an asymptotic solution of order 1 of the system (5.1) if

$$F(x,u,Du,D^2u) \equiv \omega^{-1}\mathcal{R}(x,\omega) \tag{5.3}$$

with $|\mathcal{R}|$ bounded on V, uniformly in $\omega > 0$.

Remark 5.1 *If one were to write the system (5.1) in first-order form, it would be natural to perturb first derivatives with terms in ω^{-1}, hence the tensor $\underline{u}$ in the second-order equations by a term in ω^{-2}. This is what we did in Chapter 3 for weak progressive gravitational waves.*

We use the notations

$$\underline{f} := f(\underline{u}, D\underline{u}), \quad \delta u := u - \underline{u}, \quad \delta Du := Du - D\underline{u}, \tag{5.4}$$

$$\underline{\delta f} := \underline{f}'_u \delta u + \underline{f}'_{Du}\delta Du \tag{5.5}$$

with

$$\underline{f}'_u := \left(\frac{\partial f}{\partial u}\right)_{u=\underline{u},Du=D\underline{u}}, \quad \underline{f}'_{Du} := \left(\frac{\partial f}{\partial Du}\right)_{u=\underline{u},Du=D\underline{u}} \tag{5.6}$$

and, with analogous definitions,

$$\underline{\delta}^2\underline{f} \equiv \underline{f}''_{uu}(\delta u,\delta u) + 2\underline{f}''_{uDu}(\delta u,\delta Du) + \underline{f}''_{DuDu}(\delta Du,\delta Du). \tag{5.7}$$

A function $f(u,Du)$ smooth in a neighbourhood of $(\underline{u},D\underline{u})$ admits the Taylor expansion

$$f(u,Du) = \underline{f} + \underline{\delta f} + \frac{1}{2}\underline{\delta^2 f} + \mathcal{S} \tag{5.8}$$

[22] In the case of quasidiagonal systems the polarization condition is trivial; there is no need to add an ω^{-2} perturbation to satisfy the system at order zero.

The remainder $\mathcal{S}$ can be written

$$\mathcal{S} \equiv \int_0^1 f_t'''(\delta u, \delta Du)dt \tag{5.9}$$

where f_t''' denotes the third derivative of f at the point $(\underline{u} + t\delta u, D\underline{u} + tD\delta u)$, hence $f_t'''(\delta u, \delta Du)$ is generically a homogeneous cubic polynomial in δu and δDu.

We assume that the coefficients $g^{\alpha\beta}$ and f^A admit such expansions in Taylor series around the background $\underline{u}$.

With the assumed expansion (5.2) of u we have

$$u - \underline{u} = \omega^{-1}v, \quad D_\alpha u - D_\alpha \underline{u} = \phi_\alpha v' + \omega^{-1}\underline{D}_\alpha v, \tag{5.10}$$

where we denote by $\underline{D}$ a covariant derivative, in a given metric $\underline{g}$, of a tensor on V depending on the parameter ξ:

$$(\underline{D}_\alpha v)(x, \omega\phi(x)) := \{D_\alpha v(x, \xi)\}_{\xi=\omega\phi(x)} \tag{5.11}$$

and for the second derivative

$$D^2_{\alpha\beta}u = D^2_{\alpha\beta}\underline{u} + \omega\phi_\alpha\phi_\beta v'' + \phi_\beta \underline{D}_\alpha v' + \phi_\alpha \underline{D}_\beta v' + \omega^{-1}\underline{D}^2_{\alpha\beta}v. \tag{5.12}$$

The expansion (5.8) applied to $g^{\alpha\beta}$ and f^A shows that the principal term in the asymptotic expansion with respect to ω of the left-hand side, F^A, of Equation (5.1) is in ω and contains a product $v'v''$ if $g^{\alpha\beta}$ depends on Du. Since this dependence on Du in the coefficient $g^{\alpha\beta}$ does not occur in the physical fields which interest us, we consider in the following only hyperquasilinear equations, i.e. the case where $g^{\alpha\beta}$ depends only on u.

We remark now that if the third derivative $\partial^3 f/\partial(Du)^3$ is not zero at $(\underline{u}, \underline{Du})$, then a term of order zero in ω appears in the remainder of the asymptotic expansion of f^A. We consider in the following equations where f^A is quadratic in Du, as occurs in physical fields of interest to us here.

5.2 *Hyperquasilinear systems with f quadratic in Du*

We consider quasidiagonal-hyper quasilinear second-order systems of the type with g a Lorentzian metric in the neighbourhood of a given field $\underline{u}$, q a quadratic form in Du with coefficients function of u, and a a linear form,

$$F(u, Du, D^2u) \equiv g^{\alpha\beta}(u)D^2_{\alpha\beta}u + q^{\alpha\beta}(u)(D_\alpha u, D_\beta u) + a^\alpha(u)D_\alpha u + b(u) = 0;$$

that is, in coordinates,

$$g^{\alpha\beta}(u)D^2_{\alpha\beta}u^A + q^{A,\alpha\beta}_{BC}(u)D_\alpha u^B D_\beta u^C + a^{A,\alpha}_B(u)D_\alpha u^B + b^A(u) = 0. \tag{5.13}$$

Inserting expressions (5.2) of u and DU in the Taylor expansions of $g^{\alpha\beta}$ and f, one finds

$$F(u, Du, D^2u)(x) \equiv \{\omega F^{(-1)} + F^{(0)} + \omega^{-1}\mathcal{R}\}_{\xi=\omega\phi(x)}$$

with $F^{(-1)}$, $F^{(0)}$ and $\mathcal{R}$ bounded on an open set $W \subset V$, uniformly in ξ if it is so of v and its derivatives with respect to x and ξ.

5.2.1 *Phase*

The field u is an asymptotic solution of order -1 if $F^{(-1)} = 0$. We have

$$F^{(-1)}(x,\xi) \equiv \underline{g}^{\alpha\beta}\varphi_\alpha\varphi_\beta v''(x,\xi); \tag{5.14}$$

therefore $F^{(-1)} = 0$ implies for a non-zero wave that the phase ϕ is isotropic for the background,

$$\underline{g}^{\alpha\beta}\phi_\alpha\phi_\beta = 0. \tag{5.15}$$

There is no polarization condition on v.

5.2.2 *Propagation and back reaction*

The field u is an asymptotic solution of order zero only if, in addition to (5.15), it holds that $F^{(0)} = 0$. Straightforward computation shows that $F^{(0)} - \underline{F}$ is the sum of a linear term in v' and a non-linear term,

$$F^{(0)}(x,\xi) \equiv \underline{F}(x) + \mathcal{P} + \mathcal{N}$$

where the linear term is a propagation operator for $v\prime$ along the rays of the phase

$$\mathcal{P}(v') := \underline{g}^{a\beta}\{2\varphi_\alpha D_\beta v' + v' D_\alpha\varphi_\beta\} + 2\underline{q}^{\alpha\beta}\varphi_\beta D_\alpha\underline{u}v' + \underline{a}^\alpha\varphi_\alpha v' \tag{5.16}$$

and the non-linear term $\mathcal{N}$ is

$$\mathcal{N} := \underline{q}^{\alpha\beta}\varphi_\alpha\varphi_\beta v'^2 + \underline{g'}_u^{\alpha\beta}\varphi_\alpha\varphi_\beta v v''. \tag{5.17}$$

The equation $F^{(0)} = 0$ is a quasilinear first-order partial differential equation for the field $(x,\xi) \mapsto v'(x,\xi)$. It reduces to an ordinary differential equation on V for v' if the metric g satisfies the **exceptionality condition**

$$\underline{g'}_u^{\alpha\beta}\phi_\alpha\phi_\beta = 0. \tag{5.18}$$

This condition is satisfied by semilinear equations, and by the quasidiagonal system of the Einstein equations in harmonic gauge.

We deduce from the formulas (5.16), (5.17) the following theorem.

Theorem 5.2 *A progressive wave on the Lorentzian manifold* $(V, \underline{g} := g(\underline{u}))$

$$u(x) = \underline{u}(x) + \omega^{-1}\{v(x,\xi)\}_{\xi=\omega\phi(x)} \tag{5.19}$$

is an asymptotic solution of order 1 on $V \times R$ *for 5.13 if the following conditions are satisfied:*

1. *The phase* ϕ *satisfies the eikonal equation of the background* $\underline{g}$.
2. *The metric* g *satisfies at* $\underline{u}$ *the exceptionality condition*

$$\underline{g'}_u^{\alpha\beta}\phi_\alpha\phi_\beta = 0, \tag{5.20}$$

3. v' *satisfies on* $V \times R$ *the ordinary semilinear differential equation*

$$\mathcal{P}(v') + \underline{q}^{\alpha\beta}(v', v')\phi_\alpha\phi_\beta + \underline{F}(x) = 0 \tag{5.21}$$

4. $v(x, \xi)$ *is bounded for* $x \in V$*, uniformly in* ξ*.*

We complete this theorem by the following lemma.

Lemma 5.3 *A necessary condition for the boundedness of* v *is that* v' *and the background* $\underline{u}$ *satisfy the "**back reaction**" equation*

$$\underline{F}(x) = \lim_{T=\infty} \frac{1}{T} \int_0^T -\underline{q}^{\alpha\beta}(v'(x,\xi), v'(x,\xi))\varphi_\alpha\varphi_\beta d\xi. \tag{5.22}$$

Proof Since $\mathcal{P}$ is linear and homogeneous in v' we have if v is bounded

$$\lim_{T=\infty} \frac{1}{T} \int_0^T \mathcal{P}(v')d\xi = \mathcal{P}\left(\lim_{T=\infty} \frac{1}{T} \int_0^T v' d\xi\right) = 0. \tag{5.23}$$

The condition (5.22) results from averaging Equation (5.21). □

Remark 5.4 *If* $\underline{F(x)} = 0$ *a sufficient condition for (5.22) to be satisfied is that*

$$\underline{q}^{\alpha\beta}(v'(x,\xi)v'(x,\xi))\varphi_\alpha\varphi_\beta := \underline{q}^{\alpha\beta}_{AB} v'^{A\prime} v'^{B} \varphi_\alpha\varphi_\beta = 0;$$

this condition is a fortiori satisfied if for all A, B

$$\underline{q}^{\alpha\beta}_{AB}\varphi_\alpha\varphi_\beta = 0. \tag{5.24}$$

5.3 *The null condition*

Conditions (5.20) and (5.24) are a generalization of the **null condition** introduced independently by Christodoulou and Klainerman[23] to prove the global existence of solutions of non-linear wave equations on Minkowski spacetime (see Chapter 15). It has also been used by Klainerman and collaborators to lower the regularity demanded of solutions.

Definition 5.5 *A quasilinear quasidiagonal second-order differential system on a Minkowski spacetime* M^{n+1}*,*

$$F(u, Du, D^2u) = 0$$

which admits $\underline{u} \equiv 0$ *as a solution, i.e.* $\underline{F} \equiv F(0,0,0) = 0$*, is said to satisfy the **null condition** if*

1. *The linearisation at* $\underline{u} = 0$ *of the system is the wave equation in the Minkowski metric:*

$$\underline{\delta F} = \eta^{\alpha\beta} D^2_{\alpha\beta} \delta u.$$

[23] Christodoulou, D. (1986) *Commun. Pure Appl. Math.*, **39**, 267; Klainerman, S. (1986) *Lectures in Applied Maths*, **23**.

2. *The second derivative at $u = 0$ is such that:*

$$\underline{\delta}^2 \underline{F} \equiv \underline{F}''(\delta u, \delta Du, \delta D^2 u) \equiv 0$$

whenever $\delta u, \delta Du, \delta D^2 u$ are replaced by the following tensors:

$$\delta u = X, \quad \delta Du = Y \otimes \ell, \quad \delta D^2 u = Z \otimes \ell \otimes \ell$$

with X, Y, and Z arbitrary tensors and ℓ a covector null for the Minkowski metric.

We say that a system of the type (5.1) satisfies a **generalized null condition**[24] on a background $\underline{u}$ if it satisfies conditions (5.20) and (5.24).

The wave map equation is an example of a second-order quasidiagonal semilinear system which satisfies the generalized null condition.

As a consequence of this definition we have the following corollary of Theorem (5.2).

Corollary 5.6 *The progressive wave*

$$u(x) = \underline{u}(x) + \omega^{-1}\{v(x, \xi)\}_{\xi = \omega\phi(x)} \tag{5.25}$$

is an asymptotic solution of the system (5.1) in a domain $W \subset V$ spanned by rays of the isotropic phase ϕ, if the system at $\underline{u}$ satisfies the generalized null condition, the background $\underline{u}$ is a solution of the system and v' satisfies the linear propagation equations along the rays of ϕ, $\mathcal{P}(v') = 0$.

6 Non quasidiagonal second-order systems

We consider an hyperquasilinear second-order system for a field u, with principal part sum of a diagonal part and a non-diagonal one which, inspired by classical field equations, we call the gauge part. We write the system

$$F(u, Du, D^2 u) \equiv G(u).D^2 u + f(u, Du) = 0 \tag{6.1}$$

with

$$G(u).D^2 u \equiv g^{\alpha\beta}(u) D^2_{\alpha\beta} u + P(u).D^2 u; \tag{6.2}$$

that is, in index notation

$$F^A(u, Du, D^2 u) \equiv g^{\alpha\beta}(u) D^2_{\alpha\beta} u^A + P_B^{A\alpha\beta}(u) D^2_{\alpha\beta} u^B + f^A(u, Du) = 0. \tag{6.3}$$

We consider the case, as in the previous section, where f is quadratic in Du. We use the same notations and we make the same hypotheses on g and f.

It is convenient in the case of non-quasidiagonal systems to introduce second-order perturbations in order to find propagation equations satisfied by the first-order perturbation, as we did for first-order systems, and if possible eliminate gauge terms from them. A progressive wave will be of the form

$$u(x) = \underline{u}(x) + \{\omega^{-1} v(x, \xi) + \omega^{-2} w(x, \xi)\}_{\xi = \omega\phi(x)}. \tag{6.4}$$

[24] Choquet-Bruhat, Y. (2000) *Ann. Phys. (Leipzig)*, **9**(3–5), 258–66.

It will be an asymptotic solution of order 1 of the system (6.1) on V if its insertion in the system gives (with the derivation law (3.7))

$$F(u, Du, D^2u)(x) = \omega^{-1}\mathcal{R}(x, \omega), \tag{6.5}$$

with $|\mathcal{R}|$ bounded on V, uniformly in $\omega > 0$. Necessary conditions are the annulation of the coefficients F^{-1} and $F^{(0)}$ in the expansion of F resulting from the insertion of u.

6.1 *Phase and polarization*

We find that

$$F^{(-1)}(x, \xi) \equiv \underline{g}^{\alpha\beta}\phi_\alpha\phi_\beta v''^A + P_B^{A,\alpha\beta}(\underline{u})\phi_\alpha\phi_\beta v''^B = 0. \tag{6.6}$$

We choose for $D\phi$ a null vector for the background metric $\underline{g} = g(\underline{u})$; then the equation above reduces to a linear homogeneous system for v'', and hence for v itself if we disregard the addition of terms independent of ξ; these equations, called **polarization conditions**, are

$$P_B^{A,\alpha\beta}(\underline{u})\phi_\alpha\phi_\beta v^B = 0. \tag{6.7}$$

The general solution is

$$v^A(x, \xi) = \sum_{1\le i\le N} h^A_{(i)}(x)U_{(i)}(x, \xi) \tag{6.8}$$

where the $h_{(i)}$ are a basis of the kernel of the linear operator $P^{\alpha\beta}(\underline{u})\phi_\alpha\phi_\beta$ and the $U_{(i)}$ are functions on $V \times R$, at this stage arbitrary.

We will use the following definition.

Definition 6.1 *A tensor v is said to be polarized (with respect to $\underline{u}$ and the phase ϕ) if*

$$\underline{P}.v \otimes D\phi \otimes D\phi = 0, \quad \textit{i.e.} \quad P_B^{A,\alpha\beta}(\underline{u})v^B\phi_\alpha\phi_\beta = 0. \tag{6.9}$$

6.2 *Propagation equations*

The equation $F^{(0)}(x, \xi) = 0$ is found to be, when ϕ_α is a null vector for $\underline{g}$,

$$F^{(0)} \equiv \underline{F}(x) + \mathcal{P} + \mathcal{N} + \mathcal{G} = 0, \tag{6.10}$$

where $\mathcal{P}$ and $\mathcal{N}$ are given by (5.16), (5.17) and the gauge term $\mathcal{G}$ is

$$\mathcal{G}^A \equiv \left[\underline{P}_B^{A\alpha\beta}w''^B + \frac{\partial P_B^{A\alpha\beta}}{\partial u^C}(\underline{u})v^C v''^B\right]\phi_\alpha\phi_\beta + 2\underline{P}_B^{a\beta,A}\left\{\phi_\alpha D_\beta v'^B + \frac{1}{2}v'^B D_\alpha\phi_\beta\right\}.$$

Looking at linear and non-linear terms in v, w and derivatives with respect to ξ, we see that solutions of the system (6.10) can be bounded for all ξ only if

$$\lim_{T=\infty}\frac{1}{T}\int_0^T \{\phi_\alpha\phi_\beta[\underline{G'_u}^{\alpha\beta}vv'' + \underline{q}^{\alpha\beta}(v', v')](x, \xi) + \underline{F}(x)\}d\xi = 0 \tag{6.11}$$

with, in index notation,

$$(\underline{G'_u}^{\alpha\beta} vv'')^A = \underline{g'^{\alpha\beta}_{u^C}} v^C v''^A + \underline{P'^{A\ \alpha\beta}_{B,u^C}} v^C v''^B. \tag{6.12}$$

Sufficient conditions (not necessary, see Section 8) for Equation (6.11) to be satisfied for polarized v are:

1. $\underline{u}$ is a solution of the system: $\underline{F} = 0$.
2. **Polarized exceptionality** condition.

$$\phi_\alpha \phi_\beta \underline{G'_u}^{\alpha\beta} vv'' = 0 \quad \text{for all polarized } v. \tag{6.13}$$

   This condition is identically satisfied by semilinear systems, since then G does not depend on u.

3. **Polarized null** condition

$$\phi^\alpha \phi^\beta \underline{q}^{\alpha\beta}(v', v') = 0 \quad \text{for all polarized } v. \tag{6.14}$$

If we assume in addition of these three conditions that v' satisfies the **Propagation equation**

$$\mathcal{P}(v') = 0,$$

then Equations (6.10) reduce to a system of linear equations for w'' with right-hand side linear in v',

$$\underline{P}_\phi . w'' = L.v', \quad \text{i.e.} \quad \underline{P}_B^{A,\alpha\beta} \phi_\alpha \phi_\beta w''^B = 2\underline{P}_B^{a\beta,A} \left\{ \phi_\alpha D_\beta v'^B + \frac{1}{2} v'^B D_\alpha \phi_\beta \right\}.$$

A necessary and sufficient condition for the existence of w'' is that $L.v'$ be orthogonal to the kernel of the dual of the linear (algebraic) operator $\underline{P}_\phi$, We will not discuss the general case, but look at applications to some field equations of physics. It turns out in these cases that this dual operator ${}^*\underline{P}_\phi$ is injective, and that the polarization of the field v is conserved by the propagation equation for v for the corresponding w.

7 Yang–Mills–scalar equations

The Yang–Mills-scalar equations are a good example to apply the construction of progressive waves. They are part of the standard model[25], of fundamental importance in physics.

7.1 *Fields and equations*

7.1.1 *Fields*

The **Yang–Mills field** on a spacetime V with given Lorentzian metric g is a connection 1-form A with values in a Lie algebra $\mathcal{G}$ of $N \times N$ matrices. The curvature F of this connection is

$$F_{\lambda\mu} \equiv D_\lambda A_\mu - D_\mu A_\lambda + [A_\lambda, A_\mu] \tag{7.1}$$

[25] For the construction of asymptotic waves including a spinor multiplet see Choquet-Bruhat, Y. and Greco, A. (1983) *J. Math. Phys.*, **24**(2), 377.

where we denote by D the covariant[26] derivative in the metric g and by $[.,.]$ the Lie bracket in the Lie algebra $\mathcal{G}$. We denote by $\hat{D}$ the g metric and A connection derivative, and raise indices with g; then

$$\hat{D}_\lambda F^{\lambda\mu} := D_\lambda F^{\lambda\mu} + [A_\lambda, F^{\lambda\mu}]. \tag{7.2}$$

The **scalar multiplet** Φ is a mapping[27] $V \to C^N$ such that

$$\hat{D}_\lambda \Phi \equiv D_\lambda \Phi + A_\lambda \Phi. \tag{7.3}$$

7.1.2 *Equations*

The equations are the **Yang–Mills equations** with source Φ

$$Y^\mu := \hat{D}_\lambda F^{\lambda\mu} - J^\mu = 0, \quad J^\mu \equiv \mathrm{Re}\, \Phi^* \hat{D}^\mu \Phi \tag{7.4}$$

and the quasidiagonal second-order differential **scalar multiplet equations**

$$SM := \hat{D}^\lambda \hat{D}_\lambda \Phi - K(\Phi) = 0, \tag{7.5}$$

where the definition of $\hat{D}$ gives

$$\hat{D}^\lambda \hat{D}_\lambda \Phi \equiv D^\lambda D_\lambda \Phi + 2A_\lambda D^\lambda \Phi + (D^\lambda A_\lambda + A^\lambda A_\lambda)\Phi. \tag{7.6}$$

Equations (7.4) are second-order differential equations for A; they are:

$$Y_\mu := D^\lambda D_\lambda A_\mu - D^\lambda D_\mu A_\lambda + D^\lambda [A_\lambda, A_\mu] + [A^\lambda, F_{\lambda\mu}] - \mathrm{Re}\, \Phi^* \hat{D}^\mu \Phi = 0 \tag{7.7}$$

Equations (7.4) and (7.5) form a system of the type studied in the previous section for the unknown $u := (A, \Phi)$. The system is semilinear and non-quasidiagonal. The gauge part contains only A and reads

$$-D^\lambda D_\mu A_\lambda. \tag{7.8}$$

7.2 *Phase and polarization*

We look for an asymptotic wave

$$u(x) \equiv \underline{u}(x) + \omega^{(-1)}\{u^{(1)}(x,\xi) + \omega^{-1} u^{(2)}(x,\xi)\}_{\xi=\omega\phi}, \quad u := (A, \Phi). \tag{7.9}$$

Using the results of the previous section, we choose a phase ϕ isotropic for the metric g. We find that u is an asymptotic solution of order -1 of the system if it satisfies the polarization condition

$$\phi_\lambda A^{(1)\lambda} = 0 \tag{7.10}$$

[26] Note that F does not depend on the metric g.

[27] More general representation spaces can be considered, but they will make notations heavier without changing the essential results.

7.3 *Propagation*

7.3.1 *General equations*

According to the general results u is an asymptotic solution of order 0 if it satisfies the equations $Y^{(0)\mu} = 0$ and $SM^{(0)} = 0$. These equations are found to be, by straightforward calculation, on the one hand

$$Y^{(0)\mu} := -2\phi_\lambda \phi^\mu A^{(2)\prime\prime\lambda} + \mathcal{P}_Y^\mu + \mathcal{D}_Y^\mu + \underline{Y}^\mu = 0 \tag{7.11}$$

with, using $D_\lambda \varphi_\mu = D_\mu \varphi_\lambda$, and the polarization condition (7.10)

$$\mathcal{P}_Y^\mu \equiv 2\phi^\lambda \underline{D}_\lambda A^{(1)\prime\mu} + A^{(1)\prime\mu} D_\lambda \phi^\lambda + [A^{(1)\prime\mu}, \phi_\lambda \underline{A}^\lambda] - \phi^\mu \operatorname{Re}\{\underline{\Phi}^* \Phi^{(1)'}\} \tag{7.12}$$

$$\mathcal{D}_Y^\mu \equiv -\phi^\mu \{\underline{D}_\lambda A^{(1)'\lambda} + [\underline{A}_\lambda, A^{(1)'\lambda}]\} \tag{7.13}$$

and, on the other hand, if $A^{(1)}$ is polarized,

$$2\phi^\lambda \underline{D}_\lambda \Phi^{(1)\prime} + \Phi^{(1)\prime} D_\lambda \phi^\lambda + 2\underline{A}_\lambda \varphi^\lambda \Phi^{(1)\prime} + \underline{\Phi} = 0. \tag{7.14}$$

Equations (7.12) and (7.14) are linear in derivatives with respect to ξ. The boundedness averaging *anzatz* implies therefore that the background is a solution of the equations; that is,

$$\underline{Y} = 0, \quad \underline{SM} = 0. \tag{7.15}$$

7.3.2 *Radiative coordinates*

To write propagation equations satisfies by the polarized Yang–Mills potential it is convenient to use, as in Section III.12, radiative coordinates, that is

$$\phi_0 = 1, \quad \phi_i = 0, \quad \phi^0 = 0. \tag{7.16}$$

The polarization condition on A then reads

$$A^{(1)0} = 0. \tag{7.17}$$

We can choose $A^{(2)}$ such that

$$\phi_\lambda A^{(2)\lambda\prime\prime} \equiv A^{(2)0\prime\prime} = \operatorname{Re}\{\underline{\Phi}^* \Phi^{(1)}\}'; \tag{7.18}$$

the differential equations (7.11) reduce to propagation equations along the rays for $A_i^{(1)\prime}$, not influenced by the scalar multiplet.

Remark 7.1 *The boundedness required of $A^{(2)0}$ implies the boundedness of a primitive of* $\operatorname{Re}\{\underline{\Phi}^* \Phi^{(1)}\}$*, i.e. of* $\Phi^{(1)}$*, with respect to* ξ*.*

8 Strong gravitational waves

Strong progressive gravitational waves are Lorentzian metrics of the form[28]

$$g(x) = \underline{g}(x) + \{\omega^{-1}v(x,\xi) + \omega^{-2}w(x,\xi)\}_{\xi=\omega\phi(x)}. \tag{8.1}$$

which are asymptotic solutions of order 1 of the Einstein equations. In this section we treat the vacuum case. The case of fluid sources is treated in a paper written in collaboration with A. Greco reproduced in this book.

8.1 *Einstein equations*

We have seen in Chapter 6 that the vacuum Einstein equations satisfied by a Lorentzian metric g on a smooth manifold V of dimension $n+1$ can be written intrinsically by endowing V with a smooth unphysical metric e; then, with D the covariant derivative in the metric e, one has

$$Ricci(g, Dg, D^2g) \equiv G(g).D^2g + q(g)(Dg, Dg) + g.Riemann(e), \tag{8.2}$$

where $G(g)$ is a linear operator with a diagonal and a gauge part given by

$$(G(g).D^2g)_{\alpha\beta} \equiv -\frac{1}{2}g^{\lambda\mu}D^2_{\lambda\mu}g_{\alpha\beta} + (P(g).D^2g)_{\alpha\beta}, \tag{8.3}$$

where the non-diagonal part is

$$(P(g).D^2g)_{\alpha\beta} \equiv \frac{1}{2}\{g^{\lambda\mu}D_\mu D_\alpha g_{\beta\mu} + g^{\lambda\mu}D_\mu D_\beta g_{\alpha\mu} - g^{\lambda\mu}D_\alpha D_\beta g_{\lambda\mu}\} \tag{8.4}$$

the non-principal term $q(g)(Dg, Dg)$ is a homogeneous quadratic form in Dg with coefficients depending only on g.

Both $G(g)$ and $q(g)$ are analytic in g as long as g is non-degenerate.

This Einstein operator is of the type studied in Section 6.

To write the asymptotic expansion of the Ricci tensor of the metric (8.1) we set

$$h := \delta g := g - \underline{g}, \quad \text{hence} \quad Dh = D(g - \underline{g}), \quad D^2h = D^2(g - \underline{g}). \tag{8.5}$$

We use the previous notations and the Taylor formula to write the expansion of the Ricci tensor of the metric (8.1):

$$Ricci(g) - Ricci(\underline{g}) = \underline{\delta Ricci} + \frac{1}{2}\underline{\delta^2 Ricci} + \cdots, \tag{8.6}$$

with

$$\underline{\delta Ricci} \equiv \underline{Ricci'_g}.h + \underline{Ricci'_{Dg}}.Dh + \underline{Ricci'_{D^2g}}.D^2h$$

[28] This expansion without the ω^{-2} correction term, was used by Isaacson, R. (1968) *Phys. Rev.* **166**, 1263–80. Isaacson applied the WKB original method to the linearized Einstein equations, and then looked for a solution of equations with source a stress energy tensor obtained by averaging the obtained perturbation. Progressive waves of the type (8.1) for the non-linear Einstein equations are constructed in Choquet-Bruhat, Y. (1969) *Commun. Math. Phys.*, **12**, 16–35.

and a corresponding formula, quadratic in (h, Dh, D^2h) for $\underline{\delta^2 Ricci}$; remark that $\underline{\delta^2 Ricci}$ does not contain the square of D^2h, because $Ricci(g)$ is linear in D^2g; remark also that $Ricci''_{gD^2g}$ and $Ricci''_{DgDg}$ are independent of the choice of the given metric e.

We compute the asymptotic expansion of the Ricci tensor of the metric (8.1) using the above formulae and the definition

$$h_{\alpha\beta}(x, \omega\phi(x)) := \{\omega^{-1} v_{\alpha\beta}(x,\xi) + \omega^{-2} w_{\alpha\beta}(x,\xi)\}_{\xi=\omega\phi(x)}, \tag{8.7}$$

which implies[29]

$$D_\lambda h_{\alpha\beta}(x, \omega\phi(x)) := \{\omega^{-1}[\underline{D}_\lambda v_{\alpha\beta}(x,\xi) + \phi_\lambda w'_{\alpha\beta}(x,\xi)] + \phi_\lambda v'_{\alpha\beta}(x,\xi)\}_{\xi=\omega\phi(x)} \tag{8.8}$$

$$+ \omega^{-2}[\underline{D}_\lambda v_{\alpha\beta}(x,\xi)]_{\xi=\omega\phi(x)}. \tag{8.9}$$

and a corresponding formula for the second derivative

$$\begin{aligned} D_\lambda D_\mu h_{\alpha\beta} = {} & \omega\varphi_\lambda\varphi_\mu v''_{\alpha\beta} + \varphi_\lambda \underline{D}_\mu v'_{\alpha\beta} \\ & + \varphi_\mu \underline{D}_\lambda v'_{\alpha\beta} + v'_{\alpha\beta}\underline{D}_\lambda\varphi^\lambda + \varphi_\lambda\varphi_\mu w''_{\alpha\beta} + \omega^{-1}\mathcal{R}_{\alpha\beta} \end{aligned}$$

8.2 *Phase and polarization*

The asymptotic expansion of the Ricci tensor starts with a term in ω^1 whose coefficient must vanish. We thus obtain the equations

$$Ricci'_{D^2_{\lambda\mu}g}\varphi_\lambda\varphi_\mu v'' = 0.$$

This is a linear homogeneous system for the second derivative with respect to the parameter ξ of the tensor v, which reads in coordinates

$$-\frac{1}{2}\varphi^\lambda\varphi_\lambda v''_{\alpha\beta} + \frac{1}{2}(\varphi_\alpha P_\beta(v') + \varphi_\beta P_\alpha(v')) = 0?$$

where P_α is the **polarization operator**

$$P_\alpha(v') \equiv \varphi_\lambda v'^\lambda{}_\alpha - \frac{1}{2}\varphi_\alpha v'^\lambda{}_\lambda.$$

Theorem 8.1 *If the phase is isotropic*[30] *the necessary and sufficient condition for the progressive wave to satisfy the Einstein equations at order zero in ω is that the tensor v satisfies the four "polarization conditions":*

$$P_\alpha(v) = 0.$$

The polarization conditions express the vanishing at order zero of the perturbation of the harmonicity functions $g^{\lambda\mu}\Gamma^\alpha_{\lambda\mu}$ (cf. Section III.12).

[29] As in previous sections, we denote by a prime a derivative with respect to ξ and underline partial derivatives with respect to x.

[30] This condition is necessary for the wave to be significant; see Section III.12.

8.3 *Propagation and back reaction*

Using previous results, or a direct computation, we find that the coefficient in power zero of ω in the expansion of $Ricci(g)$ is

$$R^{(0)}_{\alpha\beta} = \underline{R}_{\alpha\beta} + \mathcal{P}(v'_{\alpha\beta} + L(v', w'')_{\alpha\beta} + N_{\alpha\beta}(v, v', v'')$$

with $\mathcal{P}$ a linear propagation operator along the rays of $D\varphi$, namely,

$$(\mathcal{P}v')_{\alpha\beta} \equiv -\left\{\varphi^\lambda \underline{D}_\lambda v'_{\alpha\beta} + \frac{1}{2} v'_{\alpha\beta} \underline{D}_\lambda \varphi^\lambda\right\}$$

while L reads

$$L_{\alpha\beta} \equiv \frac{1}{2}\{\varphi_\alpha Q_\beta(v') + \varphi_\beta Q_\alpha(v')\} + \frac{1}{2}\{\varphi_\alpha P_\beta(w'') + \varphi_\beta P_\alpha(w'')\},$$

with

$$Q_\alpha(v') := \underline{D}_\lambda v'^\lambda_\alpha - \frac{1}{2}\underline{D}_\alpha v'^\lambda_\lambda.$$

The non-linear term $N(v, v', v'')$ comes from $\underline{\delta}^2 R_{\alpha\beta}$. We find that for polarized v it reduces to:

$$N_{\alpha\beta}(v, v', v'') \equiv \frac{1}{2}\varphi_\alpha\varphi_\beta \left\{v^{\lambda\mu} v''_{\lambda\mu} - \frac{1}{2} v^\lambda_\lambda v''^\mu_\mu + \frac{1}{2}\left(v'^{\lambda\mu} v'_{\lambda\mu} - \frac{1}{2} v'^\lambda_\lambda v'^\mu_\mu\right)\right\}.$$

Theorem 8.2 *The progressive wave*

$$\underline{g}_{\alpha\beta}(x) + \{\omega^{-1} v_{\alpha\beta}(x, \xi) + \omega^{-2} w_{\alpha\beta}(x, \xi)\}_{\xi=\omega\varphi(x)}$$

is an asymptotic solution of order 1 of the vacuum Einstein equations under the following hypotheses.

1. *the phase φ is isotropic for the background $\underline{g}$,*
2. *v satisfies the linear, homogeneous propagation equation*

$$\mathcal{P}(v') = 0$$

along the rays of the phase ϕ and v satisfies the polarization conditions on a hypersurface Σ transversal to these rays, assumed to span V. The field v is periodic in ξ on Σ.

3. *w is a solution of the linear system*

$$P_\alpha(w'') = Q_\alpha(v') + \frac{1}{4}\varphi_\alpha\{(v^{\lambda\mu} v'_{\lambda\mu} - v^\lambda_\lambda v'^\mu_\mu)\}' + \frac{1}{2}\varphi_\alpha(E - \underline{E})$$

4. *The background metric $\underline{g}$ satisfies the Einstein equations with source a null fluid:*

$$\underline{R}_{\alpha\beta} = \underline{E}\phi_\alpha\phi_\beta$$

where

$$\underline{E} \equiv \frac{1}{T}\int_0^T E(., \xi) d\xi, \quad \textit{with} \quad E \equiv \frac{1}{4}\left\{v'^{\lambda\mu} v'_{\lambda\mu} - \frac{1}{2} v'^\lambda_\lambda v'^\mu_\mu\right\},$$

Proof

1. If v' satisfies the propagation equations $\mathcal{P}(v') = 0$ on $V \times R$ and the polarization conditions on Σ transversal to rays which span V, then it satisfies the polarization conditions on $V \times R$ because the equation $\mathcal{P}(v') = 0$ implies the propagation of both $\varphi^\alpha v_{\alpha\beta} = 0$ and $v^\alpha_\alpha = 0$. Indeed,

$$\underline{g}^{\alpha\beta}(\mathcal{P}v')_{\alpha\beta} \equiv -\varphi^\lambda \underline{D}_\lambda v'^\lambda_\lambda + \frac{1}{2} v'^\lambda_\lambda \underline{D}_\lambda \varphi^\lambda.$$

Also

$$\varphi^\alpha(\mathcal{P}v')_{\alpha\beta} \equiv -\varphi^\lambda \underline{D}_\lambda(\varphi^\alpha v'_{\alpha\beta}) - \frac{1}{2}\varphi^\alpha v'_{\alpha\beta}\underline{D}_\lambda\varphi^\lambda$$

because if φ_α is a gradient and isotropic then

$$\varphi^\lambda \underline{D}_\lambda \varphi_\alpha = \varphi^\lambda \underline{D}_\alpha \varphi_\lambda = 0.$$

The coefficient of ω in the asymptotic expansion of $Ricci(g)$ is therefore zero.

2. The function $x \to v'(x, \xi)$ solution of the linear differential equation $\mathcal{P}(v') = 0$ with coefficients independent of ξ is determined on $V \times R$ if it is known on a submanifold transversal to rays which span V. It has period T in ξ if it is so of its data on the submanifold.

3. When v is known the equations for w'' are non-homogeneous linear equations, namely

$$L_{\alpha\beta} + N_{\alpha\beta} + \underline{R}_{\alpha\beta} = 0.$$

4. If v and w' have period T in ξ the following relation holds, because linear terms in v' or w'' integrate to zero,

$$T\underline{R}_{\alpha\beta} = -\int_0^T N_{\alpha\beta}(., \xi)d\xi$$

An elementary computation gives

$$N_{\alpha\beta} = -\varphi_\alpha\varphi_\beta \left[E - \frac{1}{2}\left(v'^{\lambda\mu,} v_{\lambda\mu} - \frac{1}{2} v^\lambda_\lambda v'^\mu_\mu \right)' \right]$$

from which follows the expression given for $\underline{R}$, hence the linear system for w'' given in the theorem. These linear equations have solutions on $V \times R$, periodic of period T in ξ, because the linear transposed homogeneous system has an empty kernel and the coefficients have period T. The tensor w can be chosen also of period T in ξ and the right-hand side has a zero integral on ξ on the interval $0 \leq \xi \leq T$.

□

Remark 8.3 *As in the case of weak gravitational waves (see Section III.12) the observable displacements due to a strong gravitational wave are governed by the Riemann tensor. The perturbation due to such a wave is one order greater*

in ω than the background Riemann tensor. The principal term is, as in Chapter 3, but now with a factor ω,

$$\rho_{\alpha\beta,\lambda\mu} = \frac{\omega}{2}(\phi_\mu\phi_\alpha v''_{\lambda\beta} - \phi_\alpha\phi_\lambda v''_{\mu\beta} - \phi_\mu\phi_\beta v''_{\lambda\alpha} + \phi_\beta\phi_\lambda v''_{\mu\alpha}).$$

8.4 *Example*

We construct in this section, as an example[31], a progressive wave which, together with its spherically symmetric background, is a model of a gravitational wave due to a non-spherically symmetric perturbation of a spherically symmetric star.

We take as background the spherically symmetric Vaidya metric, which reads in radiative coordinates

$$\underline{g} := -\left(1 - \frac{2m(u)}{r}\right) dt^2 - 2dudr + r^2(d\theta^2 + \sin^2\theta d\varphi^2).$$

When m is a constant the Vaidya metric coincides with the Schwarzschild metric.

We take for phase ϕ the isotropic function u and we look for a progressive wave

$$g(x,\omega u) := \underline{g} + \omega^{-1}v(x,\omega u), +\omega^{-2}w(x,\omega u), \quad \text{with} \quad x := (u,r,\theta,\phi)$$

$$x := (x^0 = u, x^1 = r, x^A), \quad : x^A,\ A = 2,3, := (\theta,\phi).$$

We have

$$\phi_0 = 1,\ \ \phi_i = 0, \quad \text{and} \quad \phi^0 = 0,\ \ \phi^1 = -1,\ \ \phi^A = 0.$$

We take (gauge choice) $v_{0\alpha} = 0$; the polarization conditions then read

$$v_{1i} = 0,$$

and

$$\underline{g}^{AB}v_{AB} \equiv \frac{1}{r^2}\left(v_{22} + \frac{1}{\sin^2\theta}v_{33}\right) = 0. \tag{8.10}$$

The propagation equations, $R^{(0)}_{ij} = 0$, are then found to reduce to

$$\partial_1 v_{AB} - \frac{1}{r}v_{AB} = 0.$$

These equations integrate to

$$v_{AB}(x,\xi) = r\gamma_{AB}(u,\theta,\phi,\xi).$$

We set

$$\gamma_{22} =: \alpha, \quad \gamma_{23} =: \beta,$$

The polarization conditions are then equivalent to

$$\gamma_{33} = -\alpha\sin^2\theta.$$

[31] Such an example has been considered before (see Isaacson and references therein), but with different hypotheses.

To have a progressive wave solution of order 1 of the vacuum Einstein equations it remains to satisfy the equations $R^{(0)}_{0\alpha} = 0$. The equations $R^{(0)}_{0i} = 0$ are found to be

$$R^{(0)}_{01} \equiv -\frac{1}{2} w''_{11} = 0,$$

$$R^{(0)}_{02} \equiv -\frac{1}{2} w''_{12} + \frac{1}{2r}\left(\partial_2 \alpha' + \frac{1}{\sin^2\theta}\partial_3 \beta' + 2\alpha' \frac{\cos\theta}{\sin\theta}\right) = 0,$$

$$R^{(0)}_{03} \equiv -\frac{1}{2} w''_{13} + \frac{1}{2r}\left(\partial_2 \beta' - \partial_3 \alpha' + \frac{\cos\theta}{\sin\theta}\beta'\right) = 0.$$

We see that, given α and β uniformly bounded functions of ξ as well as their primitives with respect to ξ, these linear equations in w''_{ij}, also linear in α' and β', are satisfied by functions w''_{1i} such that there exist functions w_{1i} which are uniformly bounded in ξ.

It remains to satisfy the equation $R^{(0)}_{00} = 0$, which reads

$$\frac{1}{2}\left(\underline{g}^{ij} w''_{ij} + \frac{1}{r^2}\left(\alpha^2 + \frac{\beta^2}{\sin^2\theta}\right)\right)'' = -\frac{1}{2r^2}\left(\alpha'^2 + \frac{\beta'^2}{\sin^2\theta}\right) - \frac{2}{r^2}\frac{dm}{du}. \tag{8.11}$$

This equation can be satisfied by a w uniformly bounded in ξ only if the right-hand side is also the second derivative of a function uniformly bounded in ξ. This will be the case if

$$\frac{1}{4}\left(\alpha'^2 + \frac{\beta'^2}{\sin^2\theta}\right) + \frac{dm}{du} = 0, \tag{8.12}$$

for example, assuming $\frac{dm}{du} < 0$, and setting $\mu(u) := 2(-\frac{dm}{du})^{\frac{1}{2}}$,

$$\alpha = \mu(u)\sin\xi\sin\theta, \quad \beta = \mu(u)\cos\xi\sin\theta\cos\theta. \tag{8.13}$$

One finds bounded w_{ij} satisfying all the required equations:

$$w_{12} = \frac{\mu(u)}{r}\cos\xi(\cos\theta + 2), \tag{8.14}$$

$$w_{13} = \frac{\mu(u)}{r}\sin\xi\sin^2\theta \tag{8.15}$$

and

$$w_{AB} = 0. \tag{8.16}$$

Remark 8.4 *The asymptotic metric obtained is axially symmetric (the coefficients do not depend on φ), but not spherically symmetric. It is conjectured*

and partially proved that, in agreement with physical intuition, no progressive gravitational wave with spherically symmetric background is spherically symmetric.

Exercise. *Show that the first term v in the wave falls off at infinity like* r^{-1}, *as the background metric. Show that the second term w falls off like* r^{-2}. *Compute the expansion of the Riemann tensor and compare with the results of the CK and KN studies in chapter XV.*

XII

GLOBAL HYPERBOLICITY AND CAUSALITY

1 Introduction

In the preceding chapters we have considered, for the Einstein equations, problems intrinsic and global in space but local in time. We have used, and sometimes proved, results of global differential Riemannian geometry. In this chapter we give the general properties of global Lorentzian geometry that we use in the global in time Einsteinian dynamics. Most of these general results have been known for a long time, but we prove some of them in the language we use in this book, to simplify the work of the reader.

The use of global Lorentzian geometry in solution of the Cauchy problem for general hyperbolic differential systems on manifolds was introduced by Leray in 1952[1], where he defined **global hyperbolicity** (Section 9).

Various geometric definitions related to causality (Section 6) were introduced later to study global problems in relativistic dynamics, especially by Penrose[2] and by Hawking[3]. These definitions have been used in proving singularity theorems (see Chapter 13). Geroch[4] defined the fundamental notion of a **Cauchy surface** (Section 11). The impact on physics of these definitions has been discussed, and conjectures have been formulated to remedy the possible lack of causality exhibited by solutions of the classical (non-quantized) Einstein equations. The most important of these conjectures, the strong cosmic censorship conjecture (Section 13) is still a subject of active research.

2 Global existence of Lorentzian metrics

Let V be a smooth manifold, always supposed to be Hausdorff, connected, and without boundary, unless otherwise specified. A manifold is said to be **paracompact** if it can be covered by a countable number of compact sets, hence a

[1] Leray, J. (1953) "Hyperbolic differential equations", mimeographed, IAS Princeton. See also Choquet-Bruhat, Y. (1967) in *Batelle Rencontres* (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.

[2] Penrose, R. (1968) in *Batelle Rencontres* (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.

[3] See references in Hawking, S. and Ellis, G. (1973) *The Large Scale Structure of Spacetime*, Cambridge University Press.

[4] Geroch, R. (1970) *J. Math. Phys.*, **11** 437–49.

countable number of coordinate charts[5]. We always suppose that the considered manifolds are paracompact.

Theorem 2.1 *The manifold V can be endowed with a smooth Riemannian metric if and only if it is paracompact.*

Proof Let $\{U_I\}$ be an atlas of V with local coordinates $\{x^i_I\}$ in U_I, and let ϕ_I be a partition of unity subordinate to this covering[6]. Denote by g_I the symmetric covariant 2-tensor on U_I with components in the coordinates

$$(g_I)_{ij} = \delta_{ij}. \tag{2.1}$$

then, the tensor g on V, $g := \sum_I \phi_I g_I$ is a Riemannian metric on V.

For the "only if" part of the proof see for instance CB-DM1 V A3. □

A Riemannian metric endows a manifold V with a positive definite distance, hence with the structure of a metric space. It is a **complete** Riemannian manifold if this metric space is complete (complete metric spaces are called Fréchet spaces). The **Hopf–Rinow** theorem[7] asserts the equivalence of this definition of completeness with the geodesic completeness, i.e. that all inextendible geodesics have an infinite proper length. Completeness of a Riemannian manifold is also equivalent to the compactness of closed bounded subsets.

It can be shown that any Riemannian metric on V admits a conformal metric which is complete[8].

We recall that a vector field X on V is called a **regular vector field** if it does not vanish at any point of V. We state the following known theorem.

Theorem 2.2

1. *A non-compact manifold always admits*[9] *a smooth regular vector field.*
2. *If V is compact it admits*[10] *a smooth regular vector field if its Euler–Poincaré characteristic*[11] *is zero. The condition is also necessary if V is orientable.*

[5] Recall that from any covering of a compact set by open sets one can extract a covering by a finite number of those sets.

[6] See for instance CB-DM1, IV B 1.

[7] See a proof for instance in Kobayashi and Nomizu, *Differential Geometry* 1 IV4, *loc. cit.*

[8] Nomizu, K. and Ozeki, H. (1961) *Proc. Am. Math. Soc.*, **12**, 889–92.

[9] Milnor, J. (1965) *Topology from the Differentiable Viewpoint*, University Press of Virginia, Charlotteville.

[10] Chern, S. (1944) *Ann. Math.*, **45**, 747–52. Remark that compact spacetimes are classically non-physical, due to the causality paradox.

[11] See for instance CB-DM1 IV B 3.

A consequence of this theorem is the following one, already stated in CB-DM1 V A.

Theorem 2.3

1. *If V is non-compact it always admits a smooth Lorentzian metric. If V is compact it admits a smooth Lorentzian metric if its Euler–Poincaré characteristic is zero.*

2. *The vanishing of the Euler–Poincaré characteristic is necessary if V is compact and orientable.*

Note that the Euler–Poincaré characteristic of an odd dimensional manifold is zero.

Proof

1. We suppose V paracompact, or compact and orientable with vanishing Euler–Poincaré characteristic. We endow V with a smooth Riemannian metric g^+. We denote by X the regular vector field which is known to exist and by u the unit vector collinear with X. We define a symmetric smooth covariant 2-tensor g on V by the formula:

$$g(v,w) \equiv -2\left(g^+(v,u)g^+(w,u)\right) + g^+(v,w), \quad \text{for all vector fields } v,w, \tag{2.2}$$

i.e. for the components in a local frame,

$$g_{\alpha\beta} = -2u_\alpha u_\beta + g^+_{\alpha\beta}. \tag{2.3}$$

This 2-tensor is a Lorentzian metric, as can be checked immediately by taking at a point $x \in V$ an orthonormal frame for the metric g^+, with one axis given by u; in such a frame the components of g_x are the Minkowskian ones $(-1,1,\ldots,1)$.

2. If V is compact and admits a smooth Lorentzian metric g, one defines a direction field on V as the eigendirection of g with respect to a Riemannian metric g^+; that is, at a point $x \in V$, the vector X defined up to a multiplicative factor by the relation

$$(g_{\alpha\beta} - \lambda g^+_{\alpha\beta})X^\beta = 0, \tag{2.4}$$

where the number λ, the eigenvalue of g with respect to g^+ is

$$\lambda = \underset{v}{Inf}\ \frac{g(v,v)}{g^+(v,v)}. \tag{2.5}$$

A theorem of Chern says that if V is orientable and admits a continuous direction field, it also admits a regular vector field, hence its Euler–Poincaré characteristic is zero. □

We call **Lorentzian manifold** (V, g) a smooth manifold V with a continuous Lorentzian metric g.

Remark 2.4 *To a Lorentzian manifold (V, g) with the timelike unit vector field u one associates an adapted Riemannian metric e by setting*

$$e_{\alpha\beta} := g_{\alpha\beta} + 2u_\alpha u_\beta.$$

3 Time orientation

Let (V, g) be a Lorentzian manifold. We have seen in Chapter 1 that the metric g defines at each point x two convex cones of vertex x, C_x^+ and C_x^-, spanned by the vectors v in the tangent plane to V at x such that[12] $g_x(v, v) \equiv g_{\alpha\beta}(x)v^\alpha v^\beta \leq 0$. Such vectors are called causal. They are called timelike if $g_x(v, v) < 0$. We always suppose that the metric g is continuous. The Lorentzian manifold (V, g) is said to be **time orientable** (and time oriented) if it is possible to distinguish continuously on V the fields of convex cones C_x^+, called the future, and C_x^-, called the past.

One says that the continuous field of cones C_x^+ endows the manifold V with a **causal structure**.

Theorem 3.1 *(V, g) is time orientable if and only if it admits a continuous field of timelike vectors X.*

Proof The necessity comes from the definition.

To prove the sufficiency we define C_x^+ by the property, with X a timelike vector,

$$v \in C_x^+, \quad \text{equivalent to} \quad g_x(X, v) < 0. \tag{3.1}$$

□

The following corollary generalizes the sufficient condition.

Corollary 3.2 *V can be endowed with a time-oriented continuous Lorentzian metric if it admits a continuous, nowhere vanishing, vector field.*

Proof Take a Riemannian metric g^+ on V. Denote by u the vector obtained by normalizing X in g^+. This vector is timelike and plays the role of the vector X of the theorem for the Lorentzian metric (2.3) defined in the proof of Theorem (2.3). □

Exercise. *The Mobius strip in dimension $n + 1$ is the manifold $V = W \times R^{n-1}$ where W is the slice of R^2, $-1 \leq x^1 \leq 1$, $x^0 \epsilon R$, whose points $x^1 = 1, x^0$ and $x^1 = -1, -x^0$, are identified. Endow V with a* locally flat, non time orientable Lorentzian metric Construct analogously orientable and non time orientable Lorentzian manifolds (see CB. DM1 V A).

In this chapter, and more generally in this book, a **time oriented Lorentzian manifold** (V, g) is called a **spacetime**.

[12] See Chapter 1. Recall that the metric has signature $(- + \cdots +)$.

4 Futures and pasts

We give in this section elementary definitions and easy to visualize properties relative to the causal structure of spacetimes; that is, time-oriented Lorentzian manifolds. We will give more elaborate definitions, requiring more abstract mathematics, in Section 8.

4.1 *Paths and curves*

We call a **path**, or **parametrized curve**, on a manifold V, a continuous mapping from an interval I of R with or without end points into V, $C : I \to V$ by $\lambda \mapsto C(\lambda)$. It is oriented by increasing λ. The path C is right [respectively left] differentiable if the functions which represent it in coordinate charts of V are right [respectively left] differentiable, i.e. the components $C^\alpha(\lambda + h) - C^\alpha(\lambda)/h$ [respectively $C^\alpha(\lambda) - C^\alpha(\lambda - h)/h$] tend to a limit when h tends to zero by positive values. The limit defines the right [respectively left] tangent vector to the path at the point $C(\lambda)$. A path is **piecewise differentiable** if it is both right and left differentiable and such that in any compact interval $[ab] \subset I$ the right and left tangent vectors coincide except at a finite number of points.

A **curve** is the image in V of a path $\lambda \mapsto C(\lambda)$. It is oriented by the parametrization. A curve is invariant under reparametrization preserving orientation, i.e. continuous, smooth and monotonously increasing mappings $I \to I'$. A curve can be considered as an equivalence class of paths under changes of parametrizations. A curve has a left end point if the interval I is closed on the left. If I is open on the left, $I = (a, ..$ we say that the curve has a left **end point** $C(a)$ if $C(\lambda)$ tends to $C(a)$ in the topology of V, when $\lambda > a$ tends to a. Analogous definitions hold for right end points. A curve with an end point is **extendible**, i.e. can be imbedded in a larger curve.

Remark that extendability is not defined by change of the interval I of definition of the mapping $C : I \to V$: if I is finite, we can always reparametrize it to make it infinite; for instance $I = (0, \frac{\pi}{2})$ can be reparametrized to $I' = (0, \infty)$ by setting $\lambda' = \tan \lambda$.

4.2 *Chronology and causality*

Definition 4.1 *A future-directed timelike [respectively null, or causal] path $I \to V$ by $\lambda \mapsto C(\lambda)$ is a piecewise differentiable path whose right tangent vectors are, for each $\lambda \in I$, future timelike [respectively null, or causal]. Analogous definitions hold for the past directed paths with right replaced by left.*

Timelike, null or causal curves are images in V of timelike, null, or causal paths.

Remark 4.2 *A causal curve never reduces to a point (we do not consider the zero vector as a null vector).*

Remark 4.3 *Instead of piecewise differentiable paths as a substratum, Penrose used trips, i.e. piecewise geodesics. Wald used C^1 paths. Chruściel*[13] *uses Lipschitzian, hence almost everywhere differentiable, paths, preparing the generalization of Section 8.*

Definition 4.4 *Let (V, g) be a time-oriented Lorentzian manifold.*

1. *A point y is said to be in the chronological past [respectively future] of a point x if there is a future [respectively past] directed timelike curve joining y to x; this chronology relation is denoted $y < x$ [respectively $x < y$] if y is in the past of x, then x is in the future of y.*

*The **chronological past** of a point x is the set, denoted $I_-(x)$, of points y which are in the chronological past of x. The chronological future $I_+(x)$ is defined similarly, with past replaced by future; that is,*

$$I_-(x) = \{y \in V, y < x\}, \qquad I_+(x) = \{y \in V, x < y\}. \tag{4.1}$$

We state as a lemma the following obvious property.

Lemma 4.5 *In a time-oriented Lorentzian manifold the relation $y < x$ is an order relation. It implies the following inclusion of chronological pasts*

$$I_-(y) \subset I_-(x) \quad \text{if} \quad y < x. \tag{4.2}$$

Proof By our definition of timelike curves, if $y < x$ and $x < y$ then $x = y$, also if $y < x$ and $z < y$ then $z < x$, these two properties characterize an order relation, $y < x < z$. The second one implies the inclusion (4.2). □

Analogous definitions and lemmas hold by replacing timelike by causal, i.e. by timelike or null. We state:

Definition 4.6 *In a time-oriented Lorentzian manifold the **causal future** of a point x, denoted $J_+(x)$, is the set of points which can be joined to x by a past-directed causal curve. The causal past $J_-(x)$ is defined similarly through future-directed causal curves.*

The order relation defined by causality is denoted $x \leq y$. It holds that

$$J_+(y) \subset J_+(x) \quad \text{if} \quad x \leq y. \tag{4.3}$$

The chronological future I_S^+ [respectively past I_S^-] of a subset S of V is the union of the future [respectively past] of its points. The causal future [respectively past] of a subset are defined similarly.

Definition 4.7 *A subset of V is called **achronal** [respectively **acausal**] if no pair of points in S can be joined by a timelike curve [respectively a causal curve].*

[13] P. Chruściel (2000) "Introduction to Lorentzian geometry and Einstein equations in the large", Lectures at the Levoca Summer School.

5 Causal structure of Minkowski spacetime

The $n+1$-dimensional Minkowski spacetime is the time-oriented Lorentzian manifold $M^{n+1} \equiv (R^{n+1}, \eta)$, the metric η is given in inertial coordinates by

$$\eta = -dt^2 + \sum_i^n (dx^i)^2. \tag{5.1}$$

This metric is invariant by translations in R^{n+1}. At any point $x_0 \in M^{n+1}$ the cone $C^+_{x_0}$ of future causal directions in the tangent space $T_{x_0}R^{n+1}$ is

$$X^0 \geq \left\{ \sum_i^n (X^i)^2 \right\}^{\frac{1}{2}}. \tag{5.2}$$

This field of cones endows M^{n+1} with a **causal structure**. In the case of the manifold R^{n+1} the cone $C^+_{x_0}$ can be identified with the following subset of R^{n+1}:

$$t - t_0 \geq \left\{ \sum_i^n (x^i - x^i_0)^2 \right\}^{\frac{1}{2}} > 0 \tag{5.3}$$

The causal structure is invariant by translation in R^{n+1}.

The following theorem is easy to prove.

Theorem 5.1 *In Minkowski spacetime M^{n+1} it holds that*

1. *$I^+_{x_0}$ is the set of all $x \in R^{n+1}$ such that*

$$(t - t_0)^2 - \sum_i^n (x^i - x^i_0)^2 > 0, \quad t > t_0. \tag{5.4}$$

   This is an open set.

2. *$J^+_{x_0}$ is the set of all $x \in R^{n+1}$ such that*

$$(t - t_0)^2 - \sum_i^n (x^i - x^i_0)^2 \geq 0, \quad t \geq t_0. \tag{5.5}$$

   This is the closure of $I^+_{x_0}$.

3. *The boundary $\partial J^+_{x_0} = \partial I^+_{x_0}$ is the set of null geodesics starting from x_0 :*

$$(t - t_0)^2 - \sum_i^n (x^i - x^i_0)^2 = 0, \quad t \geq t_0. \tag{5.6}$$

Proof The straight lines, geodesics of the Minkowski metric, starting from x_0 and with equation $x^\alpha - x^\alpha_0 = a^\alpha \lambda$, a^α constants, are respectively future-directed timelike or causal if

$$a^0 > \sum_i^n (a^i)^2 \quad \text{or} \quad a^0 \geq \sum_i^n (a^i)^2 \tag{5.7}$$

This proves that $I^+_{x_0}$, respectively $J^+_{x_0}$, contains the first subset, respectively the second subset.

On Minkowski spacetime it is easy to prove the reverse inclusions. Indeed, on a future-directed timelike path it holds that, on each C^1 piece,

$$\frac{dt}{d\lambda} > \left\{ \sum_i^n \left(\frac{dx^i}{d\lambda} \right)^2 \right\}^{\frac{1}{2}} . \tag{5.8}$$

Therefore, at points $x(\lambda)$ of the corresponding curve,

$$t(\lambda_{p+1}) - t(\lambda_p) > \int_{\lambda_p}^{\lambda_{p+1}} \left\{ \sum_i^n \left(\frac{dx^i}{dt} \right)^2 \right\}^{\frac{1}{2}} \tag{5.9}$$

The right-hand side is the Euclidean length of the curve of R^n parametrized by $\lambda \mapsto x^i(\lambda)$, hence greater than the Euclidean distance between the points $x^i(\lambda_{p+1})$ and $x^i(\lambda_p)$ of R^n. The inequality (5.9) holds therefore on each C^1 piece of the curve because of the triangle inequality for the distance in R^n.

An analogous proof holds for $J^+_{x_0}$. Result 3 is a consequence of 1 and 2. □

Analogous formulations with pasts are also true.

6 Causal structures on general spacetimes

On a general spacetime (V, g), i.e. a time-oriented Lorentzian manifold, the global situation is more complicated.

The study begins with the following theorem, where we construct a covering of V by so called fundamental subsets[14], which the corollary of the theorem will prove to be open subsets.

Theorem 6.1 *A spacetime (V, g) admits a covering by a family of bounded subsets of V, called **fundamental subsets**, of the form $I^+_{y,U} \cap I^-_{z,U}$ where $I^+_{y,U}$ and $I^-_{z,U}$ are the connected restrictions of I^+_y and I^-_z to some family of open sets U included in domains of charts of V.*

Proof It is sufficient to prove that each point of the domain Ω of a coordinates chart admits a neighbourhood contained in a subset of the given form. To simplify the writing of the proof we choose a zero shift[15] for the spacetime metric in the chart. In the natural frame of such local coordinates this metric reads

$$g = -N^2 dt^2 + g_{ij} dx^i dx^j . \tag{6.1}$$

The future [respectively past] timelike directions X at a point x satisfy the inequality

$$N^2(x)(X^0)^2 - g_{ij}(x) X^i X^j > 0, \quad X^0 > 0, \quad [X^0 < 0]. \tag{6.2}$$

[14] One can also use normal neighbourhoods (see next section), but their construction requires more regularity of the metric.

[15] See Chapter 1.

The spacetime metric is uniformly continuous on the relatively compact domain Ω for a topology defined by a Riemannian distance, in particular the Euclidean distance d_e. That is, given $\varepsilon > 0$ there exists a number η such that for any pair of points x, y in Ω satisfying $d_e(x, y) < \eta$ it holds that:

$$|N(x) - N(y)| < \varepsilon, \tag{6.3}$$

$$|g_{ij}(x)X^iX^j - g_{ij}(y)X^iX^j| < \varepsilon \sum_i (X^i)^2. \tag{6.4}$$

Consider a point y on the past timelike curve $x^i = x_0^i$ issued from some specified point $x_0 \in \Omega$ and such that $d_e(x_0, y) < \frac{\eta}{2}$. Then the triangle inequality shows that all points x of the subset Ω_0 defined by $d_e(x_0, x) < \frac{\eta}{2}$ are such that $d_e(x, y) < \eta$. To simplify the writing of the proof, we choose the local coordinates such that the metric coefficients at the point y are

$$N(y) = 1, \qquad g_{ij}(y) = \delta_{ij}; \tag{6.5}$$

the future timelike cone C_y^+ at y is then

$$X^0 > \left\{ \sum_i (X^i)^2 \right\}^{\frac{1}{2}}, \tag{6.6}$$

and in Ω_0 we have (we take $\varepsilon < 1$)

$$1 - \varepsilon < N(x) < 1 + \varepsilon, \tag{6.7}$$

and

$$(1 - \varepsilon) \sum_i (X^i)^2 < g_{ij}(x)X^iX^j < (1 + \varepsilon) \sum_i (X^i)^2. \tag{6.8}$$

The future cone at x is

$$N(x)X^0 > \{g_{ij}(x)X^iX^j\}^{\frac{1}{2}}; \tag{6.9}$$

the future timelike curves at x are such that

$$\frac{dt}{d\lambda} > N(x)^{-1} \left(g_{ij}(x) \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \right)^{\frac{1}{2}}. \tag{6.10}$$

The inequalities (6.8) and (6.9) show that as long as a future timelike curve lies in Ω_0 it holds on this curve that:

$$\frac{dt}{d\lambda} > \frac{1 - \varepsilon}{1 + \varepsilon} \left\{ \sum_i \left(\frac{dx^i}{d\lambda} \right)^2 \right\}^{\frac{1}{2}} \tag{6.11}$$

Consider now a future timelike curve C_y^+ issued from a point y with coordinates $y^0 > 0$, $y^i = x_0^i$. We remark that the integral over λ of $\{\sum_i (\frac{dx^i}{d\lambda})^2\}^{\frac{1}{2}}$ is the Euclidean length of the curve $\lambda \mapsto x^i(\lambda)$, and is therefore greater or equal to the Euclidean distance between its two end points. Therefore as long as this curve remains in Ω_0 (and connected), its points x, with coordinates x^i, t, remain included in the open set $C_{y,M}^+$, given by

$$t - y^0 > \frac{1-\varepsilon}{1+\varepsilon} \left\{ \sum_i (x^i - y^i)^2 \right\}^{\frac{1}{2}} . \tag{6.12}$$

The open set $C_{y,M}^+$ has, in the local coordinates, the form of a cone, invariant by translations.

On the other hand, consider a curve issued from y and such that

$$\frac{dt}{d\lambda} = 1 \quad \frac{dx^i}{d\lambda} = \alpha^i \lambda \tag{6.13}$$

with α^i constants satisfying

$$\left\{ \sum_i (\alpha^i)^2 \right\}^{\frac{1}{2}} < \frac{1-\varepsilon}{1+\varepsilon} \tag{6.14}$$

The inequalities (6.8) and (6.9) show that the tangent to such a curve at any point $x \in \Omega_0$ is future timelike, since on that curve

$$\frac{dt}{d\lambda} - N(x)^{-1} \left(g_{ij}(x) \frac{dx}{d\lambda}^i \frac{dx}{d\lambda}^j \right)^{\frac{1}{2}} \equiv 1 - N(x)^{-1} (g_{ij}(x) \alpha^i \alpha^j)^{\frac{1}{2}} > 0. \tag{6.15}$$

The considered curves fill the open subset $C_{y,m}^+$, cone with vertex y invariant by translation in local coordinates, given by

$$y^0 - t > \frac{1+\varepsilon}{1-\varepsilon} \left\{ \sum_i (x^i - y^i)^2 \right\}^{\frac{1}{2}} . \tag{6.16}$$

The subset $C_{y,m}^+$ contains the originally specified point x_0, whose coordinates are $t = 0$ and $x_0^i = y^i$, hence, since it is open, it contains also a neighbourhood ω_0^+ of the point x_0.

Exchanging future and past shows that connected past timelike curves issued from some point z in the future of x_0 lie in a convex cone $C_{z,M}^-$ with vertex z and fill another convex cone $C_{z,m}^-$, both invariant by translations and containing a neighbourhood ω_0^- of x_0.

Elementary geometry in R^{n+1} shows that both the intersections $C_{y,M}^+ \cap C_{z,M}^-$ and $C_{y,m}^+ \cap C_{z,m}^-$ are open and relatively compact. They satisfy the inclusions, with $\omega_0 := \omega_0^+ \cap \omega_0^-$

$$\omega_0 \subset C_{y,m}^+ \cap C_{z,m}^- \subset I_{y,\Omega_0}^+ \cap I_{z,\Omega_0}^- \subset C_{y,M}^+ \cap C_{z,M}^- . \tag{6.17}$$

These inclusions, valid for any point of the spacetime, prove that subsets of the form $I^+_{y,U} \cap I^-_{z,U}$ cover the manifold V. □

Corollary 6.2 *The sets I^+_x and I^-_x are open in V. The chronological future and past of an open set are the union of open sets. Therefore they are open sets.*

Proof A subset S is open if each of its points admits a neighbourhood contained in S. Let x_0 be a point in I^+_x; there exists a future timelike curve joining x and x_0. There exists, by the theorem, a neighbourhood ω_0 of x_0 contained in I^+_y with $x < y < x_0$. Then $\omega_0 \subset I^+_x$ since $I^+_y \subset I^+_x$ when $y < x$. □

Remark 6.3 *On a general spacetime $I^+_{y,U} \cap I^-_{z,U}$ does not necessarily coincide with $I^+_y \cap I^-_z$.*

Remark 6.4 *The sets J^+_x and J^-_x (closed in Minkowski spacetime) are not necessarily closed on a general Lorentzian manifold. Example: take as spacetime (V, η) the Minkowski spacetime minus a point x_0. Consider a point x on a light ray issued from x_0. Then J^-_x is the set of timelines and light rays of Minkowski spacetime arriving at x, minus the light ray arriving to x_0. Such a subset is not closed in V.*

Lemma 6.5 *In a fundamental open set the e-length of a connected causal curve is bounded.*

Proof In a compact set all continuous (properly) Riemannian metrics are equivalent. In particular, in the closure of a fundamental open set, the Euclidean metric e is equivalent to the Riemannian metric defined by

$$g^+_{\alpha\beta} := g_{\alpha\beta} + 2u_\alpha u_\beta, \tag{6.18}$$

with u a unit timelike vector. On a causal curve $\lambda \mapsto C(\lambda)$ it holds that

$$g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \leq 0, \quad \text{hence} \quad g^+_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \leq 2\left(u_\alpha \frac{dx^\alpha}{d\lambda}\right)^2. \tag{6.19}$$

Consider a causal curve, for instance future oriented, i.e. such that $\frac{dx^0}{d\lambda} > 0$. Take x^0 as a parameter on this arc and choose for u a unit tangent vector to the x^0 coordinate lines. The g^+ length s^+ of this connected causal arc is such that

$$s^+ \leq \int_{t_1}^{t_2} \sqrt{2} N dx^0. \tag{6.20}$$

The boundedness of s^+, and hence of the e-length of the considered arc, results from the boundedness of N. □

7 Geodesic coordinates, normal neighbourhoods

To prove properties resting on the local quasi Minkowskian metric behaviour of general spacetimes, it is convenient to introduce the geodesic normal local coordinates.

Recall that the canonically parametrized geodesic paths on a manifold (V, g) satisfy a first-order differential system on the tangent space TV, which reads in local coordinates

$$\frac{dx^\alpha}{d\lambda} = u^\alpha, \quad \frac{du^\alpha}{d\lambda} = -\Gamma^\alpha_{\beta\gamma}(x)u^\alpha u^\beta, \quad \alpha, \ldots = 0, \ldots n. \tag{7.1}$$

A classical theorem on ordinary differential equations says that this system admits a local solution, x^α, u^α, taking given initial values, if the coefficients are continuous, that is if the Christoffel symbols $\Gamma^\alpha_{\beta\gamma}$ are continuous functions of the coordinates x^α. However, this solution is unique only if the coefficients are Lipshitzian, hence if the metric is of class $C^{1,1}$ (differentiable with Lipshitzian differential). We make this hypothesis in this section. Then, for given initial values

$$x^\alpha(0) = x^\alpha_0, \quad u^\alpha(0) = a^\alpha \tag{7.2}$$

and small enough λ, the geodesic equations (7.1) have one and only one C^1 solution $\lambda \mapsto (x^\alpha, u^\alpha)$. In particular we set

$$x^\alpha = f^\alpha(\lambda, x_0, a), \quad x_0 = \{x^\alpha_0\}, \quad a = \{a^\alpha\}, \quad |\lambda| < \varepsilon(x_0, a), \tag{7.3}$$

Moreover there exists a neighbourhood W of zero in R^{n+1} such that the functions $a \mapsto f(\lambda, x_0, a)$ are C^1 for $a \in W$, $|\lambda| < \varepsilon(x_0, W)$.

It is straightforward to check that the functions of λ, with k a real number,

$$x^\alpha \equiv f^\alpha\left(k\lambda, x_0, \frac{a}{k}\right) \tag{7.4}$$

also satisfy the system (7.1) and take for $\lambda = 0$ the initial values x_0 while their derivatives take the values a. Therefore by restricting the domain of a it is possible to extend the definitions of the functions f up to the value 1 of the parameter λ; that is, there exists a neighbourhood W_0 of 0 in W, $\sum_\alpha |a^\alpha|^2 < \rho$, such that $\varepsilon(x_0, W_0) = 1$. Then the functions $x^\alpha = f^\alpha(1, x_0, X)$ are defined for $X \equiv \{X^\alpha\} \in W_0$. We prove for these functions the following lemma.

Lemma 7.1 *There exists a neighbourhood $U \subset W_0$ of x_0, called a normal neighbourhood, such that the functions $x^\alpha = f^\alpha(1, x_0, X)$ define in U a C^1 coordinate change.*

Proof By the definitions we have (we suppress in the notations the dependence on x_0 which is a specified point)

$$f^\alpha(\lambda, a) = f^\alpha(1, \lambda a) \tag{7.5}$$

with for $\lambda = 0$ (i.e. $X = 0$)

$$f^\alpha(0, a) = x_0^\alpha, \qquad \frac{df^\alpha(\lambda, a)}{d\lambda}(0, a) \equiv a^\alpha. \tag{7.6}$$

We have set $X = \lambda a$, hence the equality (7.5) implies that:

$$\frac{df^\alpha(\lambda, a)}{d\lambda} = \frac{\partial f^\alpha(1, X)}{\partial X^\beta} a^\beta; \tag{7.7}$$

One deduces from the definition that for $X = 0$

$$f^\alpha(1, 0) = x_0^\alpha. \tag{7.8}$$

and from (7.6) and (7.7)

$$\frac{\partial f^\alpha}{\partial X^\beta}(1, 0) = \delta^\alpha_\beta. \tag{7.9}$$

The Jacobian of the mapping $X \mapsto x$ is the identity for $X = 0$, hence invertible; the existence of U follows. □

The coordinates X^α so defined are called **geodesic coordinates**. They can be considered as defining a diffeomorphism from a neighbourhood of zero in the tangent space at x_0 onto a neighbourhood of x_0 in the manifold. This diffeomorphism is called the **exponential mapping** and denoted

$$x = \exp_{x_0} X. \tag{7.10}$$

In geodesic coordinates the geodesics issued from x_0 are represented by the straight lines

$$\lambda \mapsto X^\alpha = \lambda a^\alpha, \quad a^\alpha = \text{constant}. \tag{7.11}$$

Therefore, in geodesic coordinates, one has

$$\Gamma^\alpha_{\beta\gamma}(\lambda a) a^\beta a^\gamma = 0 \ \text{ for all } a, \tag{7.12}$$

hence, taking $\lambda = 0$,

$$\Gamma^\alpha_{\beta\gamma}(0) = 0, \tag{7.13}$$

the Christoffel symbols vanish at x_0

Remark 7.2 *In the case of a timelike geodesic the parameter λ is the length s from the origin $X = 0$ to the generic point X along the geodesic if the vector a is chosen with unit length*

$$s^2 = \lambda^2 = -g_{x_0}(X, X), \quad \text{if} \quad g_{x_0}(a, a) = -1, \tag{7.14}$$

where g_{x_0} is a constant quadratic form, the spacetime metric at the origin x_0 of the geodesic coordinates.

Remark 7.3 *The submanifolds $H_c := g_{x_0}(X, X) = \text{constant} \neq 0$, obtained by taking a constant length on the geodesics issued from one point, have these geodesics as orthogonal trajectories.*

Geodesic coordinates X^α are called **normal geodesic coordinates** if the frame at x_0 is orthonormal for the metric g_{x_0}. A neighbourhood with normal geodesic coordinates is called a **normal neighbourhood**.

The following theorems are proved using normal neighbourhoods. We formulate them for future sets; they have analogues for past sets.

Theorem 7.4 *Every point of a Lorentzian manifold (V, g), $g \in C^{1,1}$, admits an open neighbourhood U such that*

1. *The connected restriction $\tilde{\Gamma}^+_{x,U}$ to U of the set $\tilde{\Gamma}^+_x$ of light rays (i.e. null geodesics) issued from x is diffeomorphic to the restriction to a ball of the tangent space T_xV of the light cone Γ^+_x in the tangent space at x.*

2. *The connected restriction to U of the set of future timelike geodesics passing through x is the interior of $\tilde{\Gamma}^+_{x,U}$.*

Proof Consider a normal neighbourhood U of a point x_0 and the geodesic coordinates X^α. In the domain of these coordinates the geodesics issued from x_0 are represented by $X^\alpha = \lambda a^\alpha$, with a^α some constants. The future null geodesics are represented by generators of the future Minkowski light cone,

$$-(X^0)^2 + \sum_i (X^i)^2 = 0, \quad X^0 > 0. \tag{7.15}$$

A future timelike geodesic issued from x_0 is represented in the X coordinates by a straight line issued from the point $X = 0$ and interior to this cone, $X^\alpha = \lambda a^\alpha$, with a^α some constants such that

$$-(a^0)^2 + \sum_i (a^i)^2 < 0, \quad a^0 > 0. \tag{7.16}$$

Images of future timelike geodesics issued from x_0 span the interior

$$-(X^0)^2 + \sum_i (X^i)^2 < 0, \quad X^0 > 0. \tag{7.17}$$

of the future Minkowski light cone. □

The following theorem extends to general timelike curves the previous theorem, which considered only geodesics.

Theorem 7.5 *Let U be an open normal neighbourhood of some specified point x_0. A point $x \in U$ can be joined to x_0 by a past timelike curve if and only if it is interior to $\tilde{\Gamma}^+_{x_0,U}$, that is if its representative X satisfies (7.17). It can then be joined to x_0 by one and only one timelike geodesic.*

Proof

1. Consider the function defined in the coordinates X of a normal neighbourhood U by $f(X) \equiv -(X^0)^2 + \sum_i (X^i)^2$. Let $C : \lambda \mapsto C(\lambda)$ be a future timelike

curve starting from x_0, i.e. from $X_0 = 0$ in the coordinates X. The derivative of f along C is

$$\frac{df}{d\lambda} = \frac{\partial f}{\partial X^\alpha}\frac{dC^\alpha}{d\lambda}. \tag{7.18}$$

The function f vanishes at the origin $X_0 = 0$, as well as its first derivative $gradf := (\frac{\partial f}{\partial X^\alpha}) = 2(-X^0, X^i)$, but its second derivative at $X = 0$ is found to be

$$\frac{\partial^2 f}{\partial X^\alpha \partial X^\beta}(0) = 0, \quad \alpha \neq \beta, \quad \frac{\partial^2 f}{(\partial X^0)^2}(0) = -2, \quad \frac{\partial^2 f}{(\partial X^i)^2}(0) = 2. \tag{7.19}$$

Therefore the first derivative of f on the curve C vanishes at $X = 0$, i.e. for $\lambda = 0$, but its second derivative for $\lambda = 0$ is found to be, by elementary calculus

$$\frac{d^2 f}{(d\lambda)^2}(0) = -\frac{\partial^2 f}{\partial X^\alpha \partial X^\beta}(0)\frac{dC^\alpha}{d\lambda}(0)\frac{dC^\beta}{d\lambda}(0) = 2\left\{-\left(\frac{dC^0}{d\lambda}\right)^2 + \left(\frac{dC^i}{d\lambda}\right)^2\right\}(0) \tag{7.20}$$

which is negative since the curve is timelike at X_0. The function f starts decreasing on C, and hence becomes negative as soon as it leaves X_0, where it vanishes. The gradient of f at a point X is in our representation, supported by the straight line joining this point to the origin $X_0 = 0$, i.e. by a geodesic issued from X_0. Hence *gradf* is timelike when $f(X) < 0$. Using (7.18), one sees that f continues to decrease along the future timelike curve C. The curve C remains in the subset (7.17) as long as it remains in U and connected to x_0. □

Remark 7.6 *The submanifolds* $f(X) =$ constant *are obtained by constant distance from* X_0 *on the geodesics issued from* X_0*. These geodesics are their orthogonal trajectories, as remarked before.*

The following theorem is a consequence of properties of continuous functions in a ball of R^{n+1} and the fact that a Minkowskian future light cone in R^{n+1} with some vertex x_0 and another such light cone with vertex on a generator of the first one have in common only this generator, along which they are tangent. All directions in a tangent plane to a light cone along a generator, a null direction, are orthogonal to it; they are spacelike, except the null generator itself.

Theorem 7.7 *Let* U *be an open normal neighbourhood of some specified point* x_0*.*

1. *The point* $x \in U$ *can be joined to* x_0 *by a past causal curve if and only if its representative satisfies*

$$-(X^0)^2 + \sum_i (X^i)^2 \leq 0, \quad X^0 > 0. \tag{7.21}$$

2. *Let* x *be a point in* U *such that* $x \in J^+_{x_0,U}$ *but* $x \notin I^+_{x_0,U}$*. Then there is only one causal curve in* U *from* x_0 *to* x*; it is a null geodesic.*

3. *Let x, y be points in U such that $x \in J^+_{x_0,U}$ and $y \in I^+_{x,U}$, then $y \in I^+_{x_0,U}$.*
4. *The boundaries $\partial J^+_{x_0,U}$ and $\partial I^+_{x_0,U}$ are the same, they are the null geodesics generating $\tilde{\Gamma}_{x_0,U}$.*

We will deduce from these fundamental local results the following important global results which link, on a spacetime, chronological and causal pasts and futures and their boundaries.

Theorem 7.8 *Let (V, g) be a spacetime.*

1. *If $x \leq y$ and $y < z$, then $x < z$. In other words if $y \in J^+_x$, then $I^+_y \subset I^+_x$.*
2. *It always holds that (where, as usual, an overbar denotes the closure of a subset)*
$$J^+_x \subset \bar{I}^+_x. \tag{7.22}$$
3. *If $x \leq y$ and $x \not< y$ then the only causal curve between x and y is a null geodesic.*
4. *If J^+_x is closed, then:*
   (a) *$J^+_x = \bar{I}^+_x$.*
   (b) *The boundary $\partial J^+_x = \partial I^+_x = J^+_x \cap \mathcal{C} I^+_x$ is generated by null geodesics issued from x.*

Proof

1. Suppose $x \leq y$, i.e. $y \in J^+_x$ and $y < z$, i.e. $z \in I^+_y$. Consider a causal curve C from x to y and a timelike curve C' from y to z. Since a curve joining two points is a compact set we can cover $C : [ab] \to V$ by $\lambda \mapsto C(\lambda)$ with a finite number of normal neighbourhoods U_I centred at $C(\lambda_I)$, with $a = \lambda_n < \lambda_{n-1} < \ldots < \lambda_1 = b$ such that $C(\lambda_I) \subset U_{I+1}$. Considering the local theorem (7.4) in the neighbourhood U_1 centred at $C(b) = y$, we see that there is a timelike curve from any point of C_U to any point of C'_U. Let $y_1 = C(\lambda_1)$, we have $y_1 < y'_1 < z$. If $y_1 = x$, we are done. If $y_1 \geq x$ we obtain in U_2 points $y_2 \leq y_1 < y'_2$. In a finite number of steps we reach the point x and we have proved that $x < z$.
2. Suppose $y \in J^+_x$. By the local representation there exist sequences of points z_n in I^+_y converging to y. By the previous theorem $z_n \in I^+_x$, therefore $y \in \bar{I}^+_x$.
3. A curve joining two points is by definition a compact subset of V, since it is a continuous mapping from a compact interval of R into V. A causal curve C joining points x and y can therefore be covered by a finite number of open normal neighbourhoods $U, U_1, \ldots, U_{n+1}$ with centres on the curve. We set $x_0 = x \in U$, $x_{n+1} = y \in U_{n+1}$ and choose points x_p, y_p in $U_p \cap C$, $p = 1, \ldots, n$, $x_p \leq y_p$, $x_p \neq y_p$, such that also $x_{p+1} \subset U_p$ and lies on C between x_p and y_p, while y_{p+1} lies also in U_{p+2}. The local theorem shows

that the curve C is a null geodesic between any pair of points x_p, y_p, hence a null geodesic between any pair of its points.

4. (a) It holds that $I^+_x \subset J^+_x \subset \bar{I}^+_x$, hence $\bar{I}^+_x \subset \bar{J}^+_x \subset \bar{I}^+_x$, therefore $J^+_x \subset \bar{I}^+_x$ if $J^+_x = \bar{J}^+_x$.
(b) $z \in \partial I^+_x$ means that $z \in \bar{I}^+_x \cap \mathcal{C}I^+_x$, i.e. under the hypothesis, $z \in J^+_x \cap \mathcal{C}I^+_x$. It results from 3 that the curve joining x and z in ∂I^+_x is a null geodesic. □

8 Topology on a space of paths

The definition by Leray of global hyperbolicity, as well as various other global notions introduced later for Lorentzian manifolds, require the definition of limits of paths and curves; that is, the definition of a topology on a space of such objects. A satisfactory topology can be obtained for causal paths after extending their definition from piecewise C^1 paths to rectifiable paths, which we do now.

8.1 *Rectifiable paths*

Recall that a path joining two points x and y of a manifold V is a mapping $C : [a, b] \to V$ by $\lambda \mapsto C(\lambda)$ with $C(a) = x, C(b) = y$. A topology in a space of paths depends on the topology of V.

One endows a manifold V with a Fréchet[16] topology by considering a smooth properly Riemannian, complete metric e on V. Such a metric always exists. All Riemannian metrics compatible with the manifold structure are uniformly equivalent on any compact subset, the topological results are independent of the choice of e as long as we consider only a compact subset of V. We have denoted by $d_e(x, y)$ the distance in the metric e of two points $x, y \in V$: it is equal to the length of a minimizing geodesic in the metric e from x to y and satisfies the triangle inequality

$$d_e(x, y) \leq d_e(x, z) + d_e(z, x). \tag{8.1}$$

On a non-compact set a change in the metric e leads to inequivalent results. For instance if we consider the Minkowski spacetime amputated of one point, the Euclidean metric is not complete and cannot be used in some of the theorems stated below.

If a path joining two points x and y of V is C^1 by pieces its length is the finite sum of the lengths of its pieces. The length in the e metric of a C^1 piece $[a, b] \to V$ is given by the integral:

$$\ell_e(C) = \int_a^b \left\{ e\left(\frac{dC}{d\lambda}, \frac{dC}{d\lambda}\right)\right\}^{\frac{1}{2}} d\lambda \tag{8.2}$$

[16] Complete metric space.

Consider now a C^0 path $C : [a, b] \to V$ by $\lambda \mapsto C(\lambda)$. Denote by Λ_n an increasing set of numbers

$$a = \lambda_0 < \lambda_1 \leq \cdots < \lambda_n = b. \tag{8.3}$$

We denote by $\ell_e(\Lambda_n)$ the length of the "polygonal", piecewise geodesic path C_n

$$\ell_e(\Lambda_n) = \sum_{i=1}^{n} d_e(C(\lambda_i), C(\lambda_{i-1})). \tag{8.4}$$

It results from the triangle inequality that $\ell_e(\Lambda_n)$ increases when refining the subdivision of the interval, i.e.

$$\ell_e(\Lambda_n) \leq \ell_e(\Lambda_{n'}) \tag{8.5}$$

when $\Lambda_n \subset \Lambda_{n'}$. One says that the path C is rectifiable and has length $\ell_e(C)$ if $\ell_e(\Lambda_n)$ tends to a finite limit $\ell_e(C)$ when n tends to infinity. A Lipschitzian path C is rectifiable because, for such a path, by definition, there exists a number k such that, for the approximating "polygonals" Λ_n it holds that:

$$\ell_e(\Lambda_n) = \sum_{i=1}^{n} d_e(C(\lambda_i), C(\lambda_{i-1})) \leq k \sum_{i=1}^{n} (\lambda_i - \lambda_{i-1}) \leq k(b - a). \tag{8.6}$$

The length is by its definition an additive function, meaning that if $C_1 : [ab] \to V$ and $C_2 : [bc] \to V$ are rectifiable curves then $C = C_1 \cup C_2 : [ac] \to V$ has length $\ell_e(C) = \ell_e(C_1) + \ell_e(C_2)$. In particular, rectifiable paths can be parametrized by their length from their origin, or proportionally to arc length in order to have them defined as mappings from $[0, 1]$ into V.

8.2 *Topology on sets of rectifiable paths*

One denotes by $\mathcal{P}(x, y)$ the set of rectifiable paths, parametrized proportionally to their e-length, joining two points x and y of (V, e) endowed with a metric topology by defining as follows[17] the distance $\mathcal{D}$ between two such paths

$$\mathcal{D}(C_1, C_2) \equiv Sup_{\lambda \in [0,1]} d(C_1(\lambda), C_2(\lambda)) + |\ell_e(C_1) - \ell_e(C_2)| \tag{8.7}$$

In this topology of $\mathcal{P}(x, y)$ the $e-$ length is a continuous function. This topology is a Fréchet topology; that is, $\mathcal{P}(x, y)$ is a complete, infinite dimensional metric space, i.e. each Cauchy sequence of rectifiable paths admits a limit, a rectifiable path.

We denote by $\mathcal{C}(x, y)$ the subset of $\mathcal{P}(x, y)$ union of causal paths joining x and y, $x \leq y$, endowed with the induced topology, and by $\bar{\mathcal{C}}(x, y)$ its closure. An element of $\bar{\mathcal{C}}(x, y)$ is called a future causal path.

We denote by $f : [0, 1] \times \mathcal{P}(x, y) \to V$ the mapping $(\lambda, C) \mapsto C(\lambda)$. By the definition of the topology of $\mathcal{P}(x, y)$, if a rectifiable path is in the closure of $\mathcal{C}(x, y)$, each of its points is in the closure of $J_x^+ \cap J_y^-$ in V. Therefore, $J_x^+ \cap J_y^- \subset f(\bar{\mathcal{C}}(x, y)) \subset \bar{J}_x^+ \cap \bar{J}_y^-$. Hence $f(\bar{\mathcal{C}}(x, y)) = J_x^+ \cap J_y^-$ if this subset is closed in V.

[17] Bott, R. (1968) *Batelle Rencontres*, (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.

9 Global hyperbolicity

9.1 *Definition and first criterion*

Definition 9.1 *(Leray, 1952) The spacetime (V, g) is called globally hyperbolic if the set $\bar{\mathcal{C}}(x, y)$ of causal paths joining an arbitrary pair of points is a compact subset of $\mathcal{P}(x, y)$.*

Recall that a **curve** in V is the image in V of a path, mapping from an interval of R into V.

Theorem 9.2 *Global hyperbolicity forbids the presence of closed causal curves.*

Proof In a metric space a criterion of compactness of a compact subset is the possibility of extracting from every sequence a convergent subsequence. Let γ be a closed causal curve in V. Suppose it is parametrized by a parameter λ varying in an interval $[a, b] \in R$, such that $\gamma(a) = \gamma(b)$. Consider the sequence of paths $C_n : [a, a+n(b-a)] \to V$ obtained by describing γn times. Such a sequence does not admit any subsequence converging to a rectifiable path because the e length of C_n increases from a fixed quantity, $\ell_e(\gamma)$ when n increases to $n+1$. □

In fact this theorem is a corollary of the following useful criterion.

Theorem 9.3 *A necessary and sufficient condition for a spacetime to be globally hyperbolic is that in a smooth complete Riemannian metric e on V all causal paths joining two given points x and y have an e-length uniformly bounded by a number $K_{x,y}$.*

Proof

1. Necessity: The e-length of paths is a continuous function in the chosen topology of $\mathcal{P}(x, y)$. If $\bar{\mathcal{C}}(x, y)$ is a compact set this function attains on it a maximum $K_{x,y}$.

2. Sufficiency. By definition $\bar{\mathcal{C}}(x, y)$ is a closed subset of $\mathcal{P}(x, y)$, space of continuous maps from the interval [0,1] into the metric space (V, e). By the Ascoli theorem this subset is compact if:

(a) The closure of each subset of V defined by $\{C(\lambda), C \in \bar{\mathcal{C}}(x, y)\}$ is a compact subset of V.

(b) The family $C \in \bar{\mathcal{C}}(x, y)$, $C : [0, 1] \to V$ is equicontinuous in the topology of V.

Proof of (a). Since (V, e) is a complete Riemannian manifold a closed subset is compact if it is bounded. It is so for the considered subsets because by hypothesis

$$d_e(x, C(\lambda)) \leq \ell_e(C) \leq K_{x,y} \quad \text{for all} \quad C \in \bar{\mathcal{C}}(x, y). \tag{9.1}$$

Proof of (b). The path $C : \lambda \mapsto C(\lambda)$ is parametrized proportionally to its e-length from x. Therefore, the length of the path C between the points with

parameters λ_1 and λ_2 is

$$\ell_e(C(\lambda_1, \lambda_2)) = \ell_e(C)(\lambda_2 - \lambda_1). \tag{9.2}$$

Hence the e-distance between the points $C(\lambda_1)$ and $C(\lambda_2)$ is such that

$$d_e(C(\lambda_1), C(\lambda_2)) \leq \ell_e(C)(\lambda_2 - \lambda_1) \tag{9.3}$$

This inequality, and the hypothesis $\ell_e(C) \leq K_{x,y}$ proves the required equicontinuity. □

Remark 9.4 *It is sufficient to make the hypothesis on the e-length of piecewise differentiable causal curves. Indeed if C is a causal curve, i.e. a limit of piecewise differentiable causal curves C_n according to our definitions, then $\ell_e(C)$ is the limit of $\ell_e(C_n)$, hence bounded by K if it is so of $\ell_e(C_n)$.*

9.2 *Maximum of proper length*

The length in a Lorentzian metric does not define a topology on V, nor a topology in the space of paths. The proper length of a timelike curve is not a continuous function in the previously defined topology on the space of rectifiable paths (see Wald, 1984, Fig. 9.5). However, it can be proved that the length ℓ in the spacetime metric of timelike curves joining two given points of a globally hyperbolic spacetime is an upper semicontinuous function of the curve[18].

From this property there results the following theorem:

Theorem 9.5

1. *In a globally hyperbolic spacetime there exists between two points $x < y$ a timelike curve realizing the maximum of (proper) length of curves in $\mathcal{C}(x, y)$.*
2. *Such a curve is C^1 and satisfies the geodesic equation if the spacetime metric g is $C^{1,1}$.*

Proof

1. Since $\mathcal{C}(x, y)$ is a compact subset of the set of paths joining x and y, the upper semicontinuous function ℓ attains its maximum on an element γ of this set. This maximum is positive, since $x < y$, ℓ takes positive values.

2. Take on γ a point z_1 and a point z_2 in a normal neighbourhood of z_1, suppose that γ is not a geodesic between z_1 and z_2. Replace γ between these two points by the geodesic joining them. We obtain a curve of greater length, contradicting the hypothesis that γ has maximum length. □

It can also be proved that there exists a geodesic, orthogonal to M, realizing the maximum of length between a point y and points of a Cauchy surface M (see Section 11).

[18] That is, given a path $\gamma \in \mathcal{C}(x, y)$ and any number there exists neighbourhood $U(\gamma)$ such that $\ell(\tilde{\gamma}) \leq \ell(\gamma) + \varepsilon$ for all $\tilde{\gamma} \in \mathcal{C}(x, y) \cap U(\gamma)$. This property can be proved using only strong causality (see Wald, 1984, proposition 9.4.1.)

9.3 *Images in V of subsets of $\mathcal{P}(x,y)$*

We denote by f the mapping from the space $\mathcal{P}(x,y)$ of paths joining x and y into V defined by

$$f : [0,1] \times \mathcal{P}(x,y) \to V \quad \text{by} \quad (\lambda, C) \mapsto C(\lambda). \tag{9.4}$$

Theorem 9.6 *If (V,g) is globally hyperbolic then $f(\bar{\mathcal{C}}(x,y))$ is a compact subset of V for any pair (x,y), and $f(\bar{\mathcal{C}}(x,y)) = J_x^+ \cap J_y^-$.*

Proof The source of the continuous mapping $[0,1] \times \bar{\mathcal{C}}(x,y) \to V$ by $(\lambda, C) \mapsto C(\lambda)$ is a compact set if (V,g) is globally hyperbolic, the image is therefore a compact set by a general theorem of topology[19].

Suppose that $z \in f(\bar{\mathcal{C}}(x,y))$, i.e. there exists a path $\gamma \in \bar{\mathcal{C}}(x,y)$ with $z \in \gamma$, and a sequence of timelike paths $\gamma_n \in \bar{\mathcal{C}}(x,y)$, such that for all $\varepsilon > 0$ it holds that

$$D(\gamma, \gamma_n) < \varepsilon \quad \text{if} \quad n > N. \tag{9.5}$$

On the other hand, if $z \notin \bar{J}_x^+ \cap \bar{J}_y^-$ there exists an e-geodesic neighbourhood U of z in $V - \bar{\omega}$, of radius $\eta > 0$. Hence:

$$d(z, z') \geq \eta \quad \text{for all } z' \in \bar{J}_x^+ \cap \bar{J}_y^-. \tag{9.6}$$

We conclude from the inequalities (9.5) and (9.6) that $\bar{J}_x^+ \cap \bar{J}_y^- \equiv f(\bar{\mathcal{C}}(x,y))$.

To prove that $\bar{J}_x^+ \cap \bar{J}_y^- = J_x^+ \cap J_y^-$ we first consider a normal neighbourhood U centred at x. The connected restriction to U of causal curves $x^{\cap}y$ have their image in the closed subset of U, $J_{x,U}^+$. A sequence of causal curves $x^{\cap}y$ can converge in the given topology of curves only if the restriction of their images to U lies in the closed subset $J_{x,U}^+$. We now use the fact that the set $\bar{J}_x^+ \cap \bar{J}_y^-$ is compact to cover it by a finite number of normal neighbourhoods and to show that $f(\bar{\mathcal{C}}(x,y)) = J_x^+ \cap J_y^-$. □

Theorem 9.2 shows that the compactness of the images in V of the subsets $\bar{\mathcal{C}}(x,y)$ of causal paths joining two points of V is not sufficient for global hyperbolicity. We will in the next sections give sufficient conditions for global hyperbolicity using only images in V.

10 Strong and stable causalities

Strong and stable causalities have been defined by Penrose, Geroch, and Hawking, as noted in the introduction. These notions are interesting by their geometric nature, and particularly for the counterexamples they furnish to global hyperbolicity.

[19] See for instance CB-DM1 I.

10.1 *Strong causality*

Strong causality expresses that for each point x and each timelike curve leaving x there is a neighbourhood of x where the timelike curve does not reenter; that is,

Definition 10.1 *(Penrose) A spacetime is called* ***strongly causal*** *if for each point x and each neighbourhood Ω of x there is a neighbourhood $U \subset \Omega$ such that $I_x^+ \cap U_x$ is a connected subset, i.e. $I_x^+ \cap U \equiv I_{x,U}^+$.*

Strong causality prevents, in particular, the existence of closed timelike curves.

Theorem 10.2 *A spacetime is globally hyperbolic if and only if it is strongly causal and for each pair $x, y \in V$ the subset $J_x^+ \cap J_y^-$ is compact in V.*

Proof

1. Proof of the if part. We endow the spacetime with a Riemannian metric e adapted to its Lorentzian metric (Remark 2.4). Suppose that the spacetime is strongly causal, and that $J_x^+ \cap J_y^-$ is compact in V. We use Theorem 9.3 to show that the spacetime is globally hyperbolic. Since $J_x^+ \cap J_y^-$ is compact it admits a covering by a finite number of fundamental neighbourhoods U_I containing only one segment of a given causal curve C from x to y; the $e-$length of this curve is less than the sum of the e-lengths of the segments determined by the intersection of C with these neighbourhoods. On a causal curve in U_I the e-length is finite and the total e-length of C is finite since C is covered by a finite number of the $U_I's$ and, by the hypothesis of strong causality, only one segment of C is contained in one U_I.

2. Proof of the "only if" part. The strong causality of a globally hyperbolic spacetime is proved[20] by proving first that any open set in V contains a set $J_U^+ \cap J_U^-$, with U a closed ball in V, and conversely any closed ball in V contains a set of this type (one says that the sets $J_U^+ \cap J_U^-$ are a basis of neighbourhoods of V). □

10.2 *Stable causality*

A strongly causal spacetime does not possess "almost closed" causal curves. This property is not stable in the space of continuous metrics. The causal stability that we now define has some analogy with strong causality, but it is stable and easier to check analytically.

Definition 10.3 *A spacetime (V, g) is* ***stably causal*** *if there exists a continuous timelike vector field v on V such that the metric $\tilde{g} := g - v \otimes v$ (which is obviously Lorentzian) possesses no closed timelike curve.*

Theorem 10.4 *A spacetime (V, g) is stably causal if and only if there exists a differentiable function $f : V \to R$ such that its gradient ∇f is a past-directed timelike vector field on (V, g). Such a function is called a* ***time function****.*

[20] Choquet-Bruhat or Penrose (1968), independently, in *Batelle Rencontres.*

Proof Suppose there exists a time function f on (V, g) and a timelike (in the metric g) curve with future tangent vector w. Then, since ∇f is past directed,

$$g(\nabla f, w) \equiv w(f) \equiv w^\alpha \partial_\alpha f > 0. \tag{10.1}$$

Therefore f is always increasing on the considered curve, which cannot be closed. We must show that the same property holds when g is replaced by a metric of the type $\tilde{g} := g - v \otimes v$. Take $v = \nabla f$, i.e. $v_\alpha = \partial_\alpha f$. Then a straightforward computation gives

$$\tilde{g}^{\alpha\beta} = g^{\alpha\beta} + \frac{v^\alpha v\beta}{1 - v^\lambda v_\lambda}. \tag{10.2}$$

Therefore, since $v = \nabla f$ is timelike for g,

$$\tilde{g}^{\alpha\beta} \partial_\alpha f \partial_\beta f = v^\alpha v_\alpha + \frac{(v^\alpha v_\alpha)^2}{1 - v^\lambda v_\lambda} = \frac{v^\alpha v_\alpha}{1 - v^\alpha v_\alpha} < 0, \tag{10.3}$$

∇f is also timelike for $\tilde{g}$, and past directed. The previous argument shows that $\tilde{g}$ admits no closed timelike curve.

To prove the "only if" part one uses arguments analogous to the arguments which prove the existence of a Cauchy surface in globally hyperbolic spacetimes (see Hawking and Ellis, 1973). □

It can also be proved, as a corollary of the theorem, that *a stably causal spacetime is strongly causal*[21].

11 Cauchy surface

The definition of a Cauchy surface was introduced by Geroch (1970). It gives such a useful criterion for global hyperbolicity that it often replaces the original definition.

11.1 *Domain of dependence. Cauchy horizon*

Recall that a curve C is inextendible if is not a strict subset of a curve C'. If the curve has an end point $C(b)$, i.e. if it is a mapping $(a, b] \to V$, it can obviously be extended to a curve $(a, b + \varepsilon)$, which can be chosen causal if C was causal, hence "inextendible" is synonymous with "without end point". The domain of dependence of a subset S of a spacetime is roughly speaking the set of points which get all information from S. For technical reasons S is supposed closed and **achronal** – i.e. no pair of its points can be joined by a timelike curve – in the following definition.

Definition 11.1 *The future domain of dependence of a closed achronal set $S \subset (V, g)$ is the set $D^+(S)$ of points $x \in V$ such that every past-directed causal inextendible curve from x intersects S. Remark that $S \subset D^+(S) \subset J^+(S)$, and since S is achronal $D^+(S) \cap I^-(S) = \not{o}$ (the empty set).*

[21] See Wald, R. (1984) *General Relativity*, Chicago University Press.

Exchanging future and past defines $D^-(S)$, the past domain of dependence of S. The **domain of dependence** of S is

$$D(S) = D^+(S) \cup D^-(S). \tag{11.1}$$

We denote as usual by an overbar a closure of a subset.

Proposition 11.2 *$x \in \bar{D}^+(S)$ if and only if every past-directed timelike inextendible curve from x intersects S.*

Proof Denote by $D'^+(S)$ the subset of V such that every past-directed timelike inextendible curve from x intersects S. The complementary set $\mathcal{C}(D'^+(S))$ is such that there is a past-directed timelike inextendible curve from x which does not intersect S. Then there is an open neighbourhood U of x such that from points $x' \in U$ there is also a past-directed timelike inextendible curve which does not intersect S. Therefore $\mathcal{C}(D'^+(S))$ is open, and $D'^+(S)$ is closed, $\bar{D}'^+(S) = D'^+(S)$. But, obviously, $D^+(S) \subset D'^+(S)$, hence $\bar{D}^+(S) \subset D'^+(S)$.

To show the converse inclusion we consider a point $x \in D'^+(S)$, and on a timelike curve issued from x a sequence of points x_n tending to x. By Theorem 7.8 we know that every causal curve issued from x_n is part of a timelike curve issued from x, and hence intersects S. Therefore $x_n \in D^+(S)$, and consequently $x \in \bar{D}^+(S)$. □

Definition 11.3 *The future **Cauchy horizon** of a subset $S \subset (V, g)$ is defined by (S excluded):*

$$H^+(S) \equiv \partial D^+(S). \tag{11.2}$$

Analogous definitions for $D^-(S)$ and $H^-(S)$, with future and past interchanged.

Example. The future Cauchy horizon of Taub's spacetime is a sphere S^3, $t = t_+$.

11.2 *Cauchy surface*

Definition 11.4 *A **Cauchy surface** M for (V, g) is a closed achronal subset such that*

$$D(M) = V; \tag{11.3}$$

that is, every inextendible causal curve intersects M, and the curve intersects M only once if it is timelike.

The following proposition justifies the name "Cauchy surface" and its interpretation as a global instant of time.

Proposition 11.5 *A Cauchy surface in an $n+1$-dimensional spacetime is an embedded C^0 n-submanifold of V.*

Proof One says that a point $x \in S$ is in the **edge** of a closed achronal subset $S \subset V$ if there exist points $y \in I_x^+$ and $z \in I_x^-$ and a timelike curve joining y to z which does not intersect S. A Cauchy surface is without edge. The proposition is a particular case of the following one. Every closed achronal set without edge is an embedded, C^0, n-submanifold of V (see for instance Wald (1984), theorem 8.3.1). □

A fundamental criterion for global hyperbolicity is as follows.

A spacetime (V, g) is globally hyperbolic if and only if it admits a Cauchy surface.

This criterion is a consequence of points 1 and 3 of the following theorem.

Theorem 11.6 *(Geroch)*

1. *A spacetime which admits a Cauchy surface is globally hyperbolic.*
2. *Two Cauchy surfaces in a spacetime are homeomorphic.*
3. *A globally hyperbolic space time is diffeomorphic to a product $M \times R$, with each $M_t \equiv M \times t$, $t \in R$, a Cauchy surface.*

Proof

1. Let M be a Cauchy surface in the spacetime (V, g). One proves, using the definition and Remark 11.1 that (V, g) is strongly causal. For the proof that $J_x^+ \cap J_y^-$ is compact, see for instance Wald (1984), lemma 8.3.8 and theorem 8.3.10.

2. By the definition of a spacetime there exists on V a nowhere vanishing timelike vector field u. Let S and S' be two Cauchy surfaces: one defines a continuous map with continuous inverse from S to S' by associating the unique intersections with these surfaces of each timelike curve integral of the vector field u.

3. Following Geroch (1970), one introduces on V a measure μ of finite total volume. We consider (see Chruściel, 2000) a covering of V by a countable family of relatively compact sets V_n with an associated partition of unity ϕ_n; one defines a function f on V by

$$f := \sum_n \frac{\phi_n}{2^n \omega_n} \tag{11.4}$$

with

$$\omega_n := \int_V \phi_n \mu_e, \tag{11.5}$$

where the measure μ_e is the volume element of a complete Riemannian metric on V. The measure $f\mu_e$ has total volume on V

$$\omega = \int_V f\mu_e = \int_V \sum_n \frac{\phi_n \mu_e}{2^n \omega_n} = \sum_n \frac{1}{2^n} = 1. \tag{11.6}$$

One denotes

$$\omega_{\pm}(x) := \int_{J_x^{\pm}} f\mu_e. \tag{11.7}$$

It can be shown (see Geroch, 1970) that these functions are continuous if J_x^+ and J_x^- are closed. The function τ defined on V by

$$\tau := \frac{\omega_-}{\omega_+}. \tag{11.8}$$

is therefore a continuous function; hence its level sets, $\tau =$ constant are closed subsets of V. To show that these level sets are Cauchy surfaces, i.e. are achronal subsets intersected by every causal curve we remark that on a future-directed timelike curve $x < y$ implies the strict inclusion of $J_x^+ \subset J_y^+$, while the weaker order inequality $x \leq y$ on a causal curve implies only the weaker inclusion $J_x^+ \subseteq$ into J_y^+. However, in a globally hyperbolic spacetime it holds that $\bar{I}_x^+ = J_x^+$, the function τ is therefore strictly increasing on all future-directed causal curves. Causal curves will therefore intersect a level set of τ at most once. They will intersect any level set if the function τ takes a given arbitrary positive preassigned value on any inextendible causal curve if τ varies from 0 to ∞ on any such curve. One shows that such a property holds by proving that in a globally hyperbolic space time ω_- and ω_+ tend to zero on an inextendible causal curve, respectively in the past and in the future. Indeed, consider a future inextendible causal curve $[a, b) \to C$ by $\lambda \mapsto C(\lambda)$ along which ω_+ does not tend to zero, i.e. there exists $\varepsilon > 0$ and a point $x = C(\lambda_0)$ on C such that

$$\omega_+(y) > \varepsilon \quad \text{for all} \quad y = C(\lambda),\ \lambda_0 \leq \lambda < b. \tag{11.9}$$

This inequality and the expression (11.4) of the function f shows that there exists k such that

$$J^+_{C(\lambda)} \cap (\overline{\cup_{i=1}^k V_i}) \neq \emptyset \quad \text{for all } \lambda_0 \leq \lambda < b. \tag{11.10}$$

Consider an increasing sequence of values λ_n tending to b, and choose for each λ_n a point z_n in $J^+_{C(\lambda_n)} \cap (\overline{\cup_{i=1}^k V_i})$. Since these points remain in a compact set, $K := \overline{\cup_{i=1}^k V_i}$ there exists a point z limit of a subsequence of these points. A point z_n of this subsequence is in $J^+_{C(\lambda_n)}$ and hence in J_x^+, since $x = C(\lambda_0) \leq C(\lambda_n) = x_n$ implies $J^+_{C(\lambda_n)} \subset J^+_{C(\lambda_0)} = J_x^+$. The limit point z is also in J_x^+ since this set is closed. Moreover, $z_n \geq x_n$ implies that the causal curve $[\lambda_0, \lambda_n] \to C(\lambda)$ is in $J_{z_n}^-$, and the curve $[a, b) \to V$ by $\lambda \mapsto C(\lambda)$ is therefore contained in the compact set $J_x^+ \cap J_z^-$. A sequence of points on the curve has a convergence subsequence to some limit point, and strong causality implies that this limit point must be an end point of the curve. Therefore this curve is not inextendible.

We have shown that the inequality (11.9) is incompatible with the hypotheses; hence in a globally hyperbolic spacetime ω_+ tends to zero on any future inextendible causal curve. The same arguments show that ω_- tends to zero on any past-directed causal curve.

The proof is complete for a C^0 Cauchy surface. The proof of the differentiability property of τ has been written up recently[22]. □

Corollary 11.7 *On a smooth globally hyperbolic space time there exists a global time function whose level sets $f =$ constant are Cauchy surfaces.*

Proof A **time function** on (V, g) is a function $f : V \to R$ such that the vector field grad$f$ is timelike, i.e. the level sets $f =$ constant are spacelike. The corollary is an easy consequence of the theorem, when the differentiability of the Cauchy surfaces and the diffeomorphism $V \to M \times R$ are proved, because the constructed Cauchy surfaces are then spacelike. □

11.3 *Global existence of solutions of linear wave equations*

A Cauchy surface in a spacetime is a subset which is intersected along a compact set by the past [respectively the future] of any compact set. Such subsets are called by Leray[23] compact towards the past [respectively towards the future]. Leray proved future [past] global existence for solutions of generalized Cauchy problems with data on manifolds compact towards the past [future] for globally hyperbolic linear operators. The result applies in particular to linear wave equations on a globally hyperbolic spacetime with data on a Cauchy surface.

Exercise. *Prove a precise theorem, in the appropriate functional spaces, using the results of Appendix III.*

11.4 *Sufficient condition from analysis for global hyperbolicity*

The following theorem gives an easily checked sufficient condition for a sliced spacetime to be globally hyperbolic. It applies to spacetimes constructed by solution of the Cauchy problem for Einstein's equations. We first give a definition.

Definition 11.8 *A spacetime (V, g) is said to be **regularly sliced** if $V \equiv M \times I$ with $I \equiv (a, b)$ an interval of R (possibly infinite), and the spacetime metric g can be written*

$$g \equiv -N^2(\theta^0)^2 + g_{ij}\,\theta^i\theta^j, \quad \theta^0 = dt, \quad \theta^i \equiv dx^i + \beta^i dt, \tag{11.11}$$

with coefficients such that:

- *The lapse N is bounded below and above by positive numbers N_m and N_M,*

$$0 < N_m \leq N \leq N_M, \tag{11.12}$$

- *The spacelike slices $(M_t,\ g_t \equiv g_{ij}dx^i dx^j)$ are complete Riemannian manifolds with t-dependent metrics uniformly bounded below by a metric $\gamma = g_{t_0}$,*

[22] See e.g. Bär, C., Ginoux, N., and Pfäffle (2007) Wave equations on Lorentzian manifolds and quantization. *Eur. Math. Soc.*, ETH Zurich and references therein.

[23] Leray, J. (1953) Princeton Lecture Notes. See Appendix D.

i.e. there exists a number $A > 0$ such that for all tangent vectors v to M it holds that, on $M \times I$,

$$A\gamma_{ij}v^i v^j \leq g_{ij}v^i v^j, \tag{11.13}$$

- *The g_t norm of the shift vector β is uniformly bounded by a number B.*

Remark 11.9 *The hypothesis on N ensures that the parameter t measures, up to a positive factor bounded above and below, the proper time along the orthogonal trajectories to the slices M_t; the definition can obviously be reformulated using another time parameter with the same geometric impact.*

The following theorem gives a sufficient condition for global hyperbolicity.

Theorem 11.10 *(Choquet-Bruhat and Cotsakis[24]) A regularly sliced spacetime is globally hyperbolic.*

Proof M_t is a spacelike submanifold because its normal n is timelike: If n is future-directed its components in the frame θ^α are, $n^0 = N^{-1}$, $n_0 = -N$, $n^i = n_i = 0$.

Let $C : I \to V$ by $\lambda \mapsto C(\lambda)$ be a future-directed causal curve; its tangent is such that,

$$g\left(\frac{dC}{d\lambda}, n\right) \equiv -N\frac{dt}{d\lambda} < 0, \quad \text{hence} \quad \frac{dt}{d\lambda} > 0. \tag{11.14}$$

Therefore t is, on C, an increasing function of λ, and C can be reparametrized using t as it cuts each M_t, $t \in I$, at most once. We show that each M_t, $t > t_0$, $t_0 < b$ is cut by any inextendible future directed causal curve issued from a point of M_{t_0}. Indeed, suppose such a curve C is only defined for $t < T$, with $T < b$. Consider a Cauchy sequence of numbers (t_n) which converges to T and the corresponding points (c_n, t_n) of the curve C, where c_n (with components $C^i(t_n)$) are points of M. To show that when (t_n) is a Cauchy sequence converging to T, these points converge to a limit point, denoted by $c(T)$, we consider the distance d in the complete metric space (M, γ) and we have:

$$d(c_n, c_m) \leq \int_{t_n}^{t_m} \left(\gamma_{ij}\frac{dC^i}{dt}\frac{dC^j}{dt}\right)^{1/2} dt. \tag{11.15}$$

We deduce from hypothesis 2 that for all $t \in [t_0, b)$,

$$\gamma_{ij}\frac{dC^i}{dt}\frac{dC^j}{dt} \leq A^{-1}g_{ij}\frac{dC^i}{dt}\frac{dC^j}{dt}, \tag{11.16}$$

and from the causality of C and assumption 1 we obtain,

$$\left(g_{ij}\left(\frac{dC^i}{dt} + \beta^i\right)\left(\frac{dC^j}{dt} + \beta^j\right)\right)^{1/2} \leq N \leq N_M. \tag{11.17}$$

[24] Choquet-Bruhat, Y. and Cotsakis, S. (2002) *J. Geom. Phys.*, **43**(4), 345–50.

The classical subadditivity of norms implies then that,

$$\left(g_{ij}\frac{dC^i}{dt}\frac{dC^j}{dt}\right)^{1/2} \leq N + \left(g_{ij}\beta^i\beta^j\right)^{1/2} \leq N_M + B. \tag{11.18}$$

Assembling these results we find that,

$$d(c_n, c_m) \leq A^{-1}(N_M + B)(t_m - t_n). \tag{11.19}$$

This inequality shows that, if (t_n) is a Cauchy sequence of numbers converging to T, the sequence (c_n) is a Cauchy sequence in the complete Riemannian manifold (M, γ), and hence admits a limit point which we call $c(T)$. The timelike curve on $[t_0, T)$ is therefore extendible.

A similar proof gives that any inextendible past-directed causal curve arriving at a point of M_{t_0} cuts each M_t with $a < t < t_0$. □

12 Globally hyperbolic Einsteinian spacetimes

In the previous sections it was not assumed that the spacetimes are solution of the Einstein equations. In this section we will consider Einsteinian spacetimes. We will limit here our theorems to the *vacuum case*. Some results can be extended to the case of non-zero sources which satisfy hyperbolic systems. We will state the corresponding results in the relevant sections.

12.1 *Existence*

The local in time, global in space, existence theorem obtained in Chapter 6, together with the theorem just proven, gives the following theorem.

Theorem 12.1 *An initial data set $(M, \bar{g}, \bar{K})$ with $\bar{g} \in M_s$, $\bar{K} \in H_{s-1}$, $s > \frac{n}{2} + 1$, which satisfy the vacuum Einstein constraints, always admits a globally hyperbolic vacuum Einsteinian development (V, g), with M embedded in V as a Cauchy surface.*

Remark. The existence of globally hyperbolic Einsteinian developments, not necessarily regularly sliced, can be proved by gluing together local in space and time Einsteinian developments and using local geometric, causal uniqueness. This geometric local uniqueness is used in the proof of the global uniqueness we treat in the next subsection.

12.2 *Global uniqueness*

Isometric spacetimes, that is spacetimes (V, g) and (V', g') such that there exists a diffeomorphism $f : V \to V'$ with the property

$$g = f * g' \tag{12.1}$$

are physically identical.

Consider two spacetimes (V, g) and (V', g'). If there exists an isometry f from (V, g) onto (W, g'), with W a strict subset of V', then (V', g') is called an **extension** of (V, g). As is already the case for the curves that are solutions of differential equations, a uniqueness theorem for a solution of Einstein equations is meaningful only for inextendible spacetimes, also called maximal.

It is not surprising that uniqueness for a solution of the Cauchy problem holds only in the domain of dependence of the initial data set. However, due to the geometric and also non-linear nature of the Einstein equations (which implies that the domain of dependence depends on the solution), the formulation of global uniqueness will go as follows. Recall that a **development** of an initial data set $(M, \bar{g}, \bar{K})$ is a spacetime (V, g) together with an embedding of M onto a submanifold M_0 diffeomorphic to M, such that the metric $\bar{g}$ and the symmetric 2-tensor $\bar{K}$ are respectively the pullback on M of the metric induced by g on M_0 and the extrinsic curvature of M_0 as submanifold of (V, g). Two isometric developments are considered identical. A vacuum Einsteinian development is a solution of the vacuum Einstein equations. A development is called globally hyperbolic if (V, g) is globally hyperbolic with M_0 as a Cauchy surface. A development is called **maximal** if it cannot be extended to another development.

Theorem 12.2 *(Choquet-Bruhat and Geroch (1969)[25]) Every Einsteinian initial data set $(M, \bar{g}, \bar{K})$ with $\bar{g} \in M_s$, $\bar{K} \in H_{s-1}$, admits one and only one Einsteinian (vacuum) globally hyperbolic maximal development if $s > \frac{n}{2} + 1$.*

The proof is by defining a partial ordering of globally hyperbolic developments of the given initial data, using the local geometric uniqueness, and then applying the Zorn lemma to obtain maximal elements. The uniqueness is proved by identifications (the Zorn lemma again) and showing that the unique maximal element so obtained is, as required, a Hausdorff manifold (see details in Choquet-Bruhat and Geroch (1969) or in Choquet-Bruhat and York (1980)[26]).

12.3 *Examples*

The Schwarzschild spacetime, the Finkelstein–Eddington spacetimes, and the Kruskal spacetimes are all globally hyperbolic.

13 Strong cosmic censorship

The cosmic censorship conjectures, originally due to Penrose, aim at proving the deterministic character of General Relativity at the classical (non-quantum) level. We will return to these conjectures in the next chapter. We give now only the following clear-cut formulation[27] on which a consensus has been attained.

[25] Choquet-Bruhat, Y. and Geroch, R. (1969) *Commun. Math. Phys.*, **14**, 329–35.

[26] Choquet-Bruhat, Y. and York, J. (1980) in *General Relativity and Gravitation* (ed. A. Held), Plenum.

[27] Already suggested in private discussions by Geroch (1969), and formalized by Eardley and Moncrief.

Remark, however, that the statement leaves open the degree of smoothness required from extendibility.[28]

Strong cosmic censorship conjecture. *The maximal globally hyperbolic vacuum Einsteinian development of generic*[29] *initial data is inextendible if the initial data* $(\bar{g}, K)$ *is asymptotically Euclidean or compactly supported.*

Remark 13.1 *It was pointed out by P. Chruściel that the requirement that the initial manifold* $(M, \bar{g})$ *be complete is not sufficient to make the conjecture plausible; initial data for a Minkowski spacetime on an hyperboloid are then a trivial counter-example. Probably a weaker requirement than being asymptotically Euclidean could be made in the non-compact case, and in both cases the word "generic" could be made more precise.*

The conjecture may be inextendibility as an Einsteinian development, or even as a Lorentzian manifold.

Examples of non-globally hyperbolic extensions of a globally hyperbolic spacetime with initial data on S^3 (hence complete, since S^3 is a compact manifold) are provided by the Taub–Nut spacetimes (see Chapter 5), extensions of the globally hyperbolic Taub space time. However, these extensions are not a counter-example to the strong cosmic censorship conjecture, because the Taub spacetime, due to its symmetries, is not generic. It is not stable against small perturbations. A theorem proved by Moncrief and Isenberg[30] is considered in favour of the conjecture because it shows that some qualitative features of Taub–Nut spacetime imply in fact the existence of an isometry group.

Theorem 13.2 *Every analytic vacuum spacetime with a compact analytic null hypersurface* Σ *fibred by closed null geodesics has a* S^1 *isometry group whose action preserves the fibres.*

The validity of this theorem without an analyticity assumption is an interesting open problem.

Globally hyperbolic Einstein developments which are geodesically complete are inextendible, and hence trivially satisfy strong cosmic censorship. We give in the next chapter a computable criterion for causal geodesic completeness.

[28] See work of M Dafermos (2008) gr–9c 3037.013.

[29] Loosely speaking "generic" means "without special properties. In a more precise mathematical sense it can be interpreted that "generic initial data" form a dense subset of the set of possible initial data in some relevant topology

[30] Moncrief, V. and Isenberg, J. (1983) *Commun. Math. Phys.*, **86**, 485–93.

XIII

SINGULARITIES

1 Introduction

In Newtonian theory as well as in Special Relativity the spacetime manifold and metric structures are *a priori* given and smooth; singularities occurring in various fields are defined by singularities in the functions or vector fields which represent them. An example is the case of the $1/r$ singularity in the potential of a point mass[1] located at the origin in the Euclidean space arena of Newtonian gravitation. The problem is different for an Einsteinian gravitational field, which itself defines the spacetime manifold and metric, (V, g). We have supposed that the manifold is smooth (which is no restriction on its topology so long as it is C^1), and the metric is at least continuous, i.e. its components in the chart of an admissible atlas are continuous functions. Most theorems require more regularity of the metric, for example the existence and uniqueness of a local geodesic issued for a given point with a given tangent vector requires a $C^{1,1}$ metric. Since the famous "singularity theorems" of Penrose and Hawking in the 1970s, the definition taken of a singular spacetime is its future or past causal geodesic incompleteness, meaning that some of its inextendible[2] timelike or null geodesics, future or past directed, have a finite proper length or a finite canonical parameter. These theorems rest on the convergence of families of causal geodesics in a spacetime satisfying a positive (or zero) energy condition; with convergence due to the attractive, and self-attractive, character of gravitation.

Examples.

- The Schwarzschild spacetime $r > 2m$ is singular because the future ingoing causal geodesics are incomplete.
- The Kruskal spacetime is singular because the boundary $r = 0$ is attained in a finite proper time (see Chapter 4).

[1] Such a singularity finds an interpretation in the framework of distribution theory as the locally integrable function on R^3 which is a solution of the Poisson equation with source a Dirac measure at the origin. Such an interpretation does not hold for $r = 0$ in the Schwarzschild metric.

[2] In this definition we do not suppose that the spacetime itself is inextendible; an extension is considered as another spacetime. The possibility of extension of spacetimes is another problem.

- The Taub spacetime is singular because it is timelike and null geodesically incomplete (see Chapter 5).
- When a causal geodesic in the Kruskal spacetime tends to $r = 0$ the scalar invariants constructed from the Riemann curvature blow up. One says that it is a curvature singularity.
- When a causal geodesic tends to $r = 2m$ in the Schwarzschild spacetime the scalar invariants constructed with the curvature remain bounded. One says that it is a non-curvature singularity.

The existence of a non-curvature singularity is an indication that the spacetime is possibly extendible, eventually as a non-globally hyperbolic spacetime, thus violating the strong cosmic censorship conjecture. The Eddington–Finkelstein spacetimes are globally hyperbolic extensions of the Schwarzschild spacetime, We have already mentioned that the Taub–Nut spacetimes are non-globally hyperbolic extensions of the Taub spacetime.

Remark. *A curvature singularity does not imply geodesic incompleteness. The geodesic flow depends only on the $C^{1,1}$ structure of the metric. Conversely, does geodesic incompleteness imply a curvature singularity? This question is linked with the strong cosmic censorship conjecture defined in the previous chapter. We will return to it later.*

In this chapter we give in Section 2 a computable sufficient condition for the future causal completeness of a spacetime, and then a sufficient condition for its future or null incompleteness.

In Sections 3–5 we give the fundamentals of the definitions pertinent to the study of incompleteness of spacetimes by the geometric methods introduced and developed by Penrose, Hawking, and their followers. We state a sample of results and sketch some proofs, but we often send the reader back to the basic book of Hawking and Ellis, or to Chapter 8 of the excellent book by R. Wald, or more recent papers. In Sections 5 and 7 we give some elements of black hole theory and comments on Penrose's **weak cosmic censorship conjecture**, which says essentially that singularities developing from smooth initial data are hidden inside black holes. The conjecture is not easy to make mathematically precise without impoverishing its possible physical content. The conjecture is a motivation for active research.

Section 8 analyzes the study by Christodoulou of the singularities in spherically symmetric solutions of the Einstein-scalar equations.

Section 9, contributed T. Damour, is an up to date survey of his and collaborators results on the BKL conjecture.

In Section 10 we explain how the Fuchs theorem (see Appendix V) permits the analysis of some types of initial (Big Bang) singularities occurring in solutions of the Einstein equations, called **AVTD**.

2 Criteria for completeness or incompleteness

2.1 *A completeness criterion*

The following theorem[3] gives a criterion for causal geodesic completeness.

We consider a spacetime (V, g) with a regularly sliced metric (Definition 12.11.8). Recall that our definition of regularly sliced includes the boundedness above and below of the lapse N,

$$0 < N_m \leq N \leq N_M, \tag{2.1}$$

These inequalities are only gauge conditions on the choice of thetime parameter t, not a geometric restriction. In consequence, we suppose in our **future completeness** theorem the interval I of definition of the parameter t to be $(a, +\infty)$. We leave to the reader formulations with other choices of the intervals of variation of N and t.

Theorem 2.1 *Sufficient conditions for future timelike and null geodesic completeness of a regularly sliced spacetime are that* [4]

1. $|\nabla N|_{\bar{g}}$ *is bounded by a function of* t *which is integrable on* $[a, +\infty)$
2. $|K|_{\bar{g}}$ *is bounded by a function of* t *which is integrable on* $[a, +\infty)$.

Proof We have seen that in a regularly sliced spacetime, all future-directed timelike curves issued at time $t_1 > a$ can be represented by a mapping $C : [t_1, +\infty) \to V$. The proper length of such a curve in the metric g is the integral on C of the g-pseudonorm of its tangent vector; that is,

$$\ell(C) \equiv \int_{t_1}^{+\infty} \left(N^2 - g_{ij} \left(\frac{dC^i}{dt} + \beta^i \right) \left(\frac{dC^j}{dt} + \beta^j \right) \right)^{1/2} dt. \tag{2.2}$$

This length is infinite if C is uniformly timelike; that is, if the g-pseudonorms of the tangent vectors dC/dt to C are bounded away from zero by a positive constant, i.e. there exists a number $k > 0$ such that on C,

$$N^2 - g_{ij} \left(\frac{dC^i}{dt} + \beta^i \right) \left(\frac{dC^j}{dt} + \beta^j \right) \geq k^2. \tag{2.3}$$

Examples of such curves are the orthogonal trajectories of the space sections M_t. The question of timelike geodesic completeness is more delicate since these curves are not necessarily uniformly timelike.

Let u be the tangent vector to a geodesic parametrized by arc length, or by the canonical parameter in the case of a null geodesic. The components of u satisfy

[3] Choquet-Bruhat, Y. and Cotsakis, S. (2002) *J. Geom. Phys.*, **43**(4), 345–50.

[4] Klainerman and Rodniauski (2008) prove extendibility of spacetimes with τ = constant foliation and $|k|_{\bar{g}}, N^{-1}|\nabla N|_{\bar{g}}$ uniformly bounded.

in an arbitrary frame the differential equations

$$u^\alpha \nabla_\alpha u^\lambda \equiv u^\alpha \partial_\alpha u^\lambda + \omega^\lambda_{\alpha\gamma} u^\alpha u^\gamma = 0. \tag{2.4}$$

In the Cauchy adapted frame θ^α the components of u are

$$u^0 = \frac{dt}{ds}, \quad u^i = \frac{dx^i}{ds} + \beta^i \frac{dt}{ds}, \tag{2.5}$$

while the Pfaff derivatives are given by

$$\partial_i \equiv \frac{\partial}{\partial x^i}, \qquad \partial_0 \equiv \frac{\partial}{\partial t} - \beta^i \frac{\partial}{\partial x^i}. \tag{2.6}$$

It holds therefore that,

$$u^\alpha \partial_\alpha u^\lambda \equiv \frac{dt}{ds}\left[\left(\frac{\partial}{\partial t} - \beta^i \frac{\partial}{\partial x^i}\right) u^\lambda\right] + \left(\frac{dx^i}{ds} + \beta^i \frac{dt}{ds}\right) \frac{\partial}{\partial x^i} u^\lambda \equiv \frac{du^\lambda}{ds}. \tag{2.7}$$

Using $u^0 \equiv dt/ds$, we write the component $\lambda = 0$ of (2.4) in the form,

$$\frac{d}{dt}\left(\frac{dt}{ds}\right) + \frac{dt}{ds}\left(\omega^0_{00} + 2\omega^0_{0i} v^i + \omega^0_{ij} v^i v^j\right) = 0, \tag{2.8}$$

where we have set

$$v^i := \frac{u^i}{u^0} \equiv \frac{dx^i}{dt} + \beta^i. \tag{2.9}$$

If u is causal it holds that,

$$g_{ij} v^i v^j \leq N^2 \leq N_M^2. \tag{2.10}$$

We denote

$$\frac{dt}{ds} = y, \tag{2.11}$$

so (2.8) becomes

$$\frac{y'}{y} = -\left(\omega^0_{00} + 2\omega^0_{0i} v^i + \omega^0_{ij} v^i v^j\right). \tag{2.12}$$

The length (or canonical parameter extension) of the curve C is

$$\int_{t_1}^{+\infty} \frac{ds}{dt}\, dt. \tag{2.13}$$

It is infinite if ds/dt is bounded away from zero, i.e. if $y \equiv dt/ds$ is uniformly bounded.

We deduce from (2.12) that,

$$\log \frac{y(t)}{y(t_1)} = -\int_{t_1}^{t} \left(\omega^0_{00} + 2\omega^0_{0i} v^i + \omega^0_{ij} v^i v^j\right) dt \tag{2.14}$$

The integrand on the right-hand side is itself a function of t which depends on the integration of the geodesic equations. However, we can formulate *sufficient* conditions on the spacetime metric under which y is uniformly bounded using inequality (2.10). The expression for the coefficients ω given in Chapter 6 implies

$$\log \frac{y(t)}{y(t_1)} = \int_{t_1}^{t} N^{-1} \left(-\partial_0 N - 2\partial_i N v^i + K_{ij} v^i v^j\right) dt. \tag{2.15}$$

We use on the curve C the relation

$$\partial_0 N = \frac{dN}{dt} - v^i \partial_i N, \tag{2.16}$$

and estimate $\partial_i N v^i$ by using the Cauchy–Schwarz inequality and the inequality (2.10). We find that

$$\log \frac{y(t)}{y(t_1)} \leq 2 \log N_m^{-1} + N_m^{-1} \int_{t_1}^{t} (|\nabla N|_{g_t} N_M + |K|_{g_t} N_M^2) dt, \tag{2.17}$$

from which the result follows. □

Corollary 2.2 *Let P be the traceless part of K. In an expanding universe, condition 2 can be replaced by the condition that the norm $|P|_{\bar{g}}$ is integrable on $[t_1, \infty)$.*

Proof We set (recall that τ is the mean extrinsic curvature of space slices)

$$K_{ij} = P_{ij} + \frac{1}{n} \tau g_{ij}. \tag{2.18}$$

In an expanding universe $\tau < 0$. If $\tau \leq 0$, then

$$K_{ij} v^i v^j \equiv P_{ij} v^i v^j + \frac{1}{n} \tau g_{ij} v^i v^j \leq P_{ij} v^i v^j, \tag{2.19}$$

and this completes the proof. □

2.2 *An incompleteness criterion*

We now give *sufficient* conditions for the integral (2.12) to be bounded on some causal geodesic, thus proving a singularity theorem (causal geodesic incompleteness). We still suppose that the lapse N is bounded above and below by a positive number.

Theorem 2.3 *A sufficient condition for future timelike or null geodesic incompleteness of a spacetime $V \equiv (M \times (t_0, \infty)$ with a regularly sliced metric is that, on some timelike or null geodesic $C : (t_0, \infty) \to V$*

1. *$|\nabla N|_{\bar{g}}$ is integrable on C.*
2. *There exists a number $k > 0$ such that for all space vectors v*

$$K_{ij} v^i v^j \geq k. \tag{2.20}$$

Corollary 2.4 *Condition 2 can be replaced by the condition that, on C, the norm $|P|_{\bar{g}}$ of the traceless part of K is integrable while $\tau|v|^2_{\bar{g}} \geq k > 0$.*

Proof We consider a timelike or null geodesic C and we denote by z the function on this curve inverse to the function y of the previous section,

$$z = \frac{ds}{dt} = y^{-1}.$$

The formula (2.15) gives

$$\log \frac{z(t)}{z(t_0)} = \int_{t_0}^{t} N^{-1} \left(\partial_0 N + 2\partial_i N v^i - K_{ij} v^i v^j\right) dt; \tag{2.21}$$

or, equivalently

$$z(t) = z(t_1) \exp \int_{t_0}^{t} N^{-1} \left(\partial_0 N + 2\partial_i N v^i - K_{ij} v^i v^j\right) dt. \tag{2.22}$$

We use on the curve C the relation

$$\partial_0 N = \frac{dN}{dt} - v^i \partial_i N, \tag{2.23}$$

and we find

$$z(t) = z(t_0) \frac{N(t)}{N(t_0)} \exp \int_{t_0}^{t} (N^{-1} \partial_i N v^i - K_{ij} v^i v^j) dt. \tag{2.24}$$

The Schwarz inequality implies

$$|N^{-1} \partial_i N v^i| \leq |\nabla N|_{\bar{g}} |N^{-1} v|_{\bar{g}} \leq |\nabla N|_{\bar{g}};$$

because if u is causal then, by (2.9), $|N^{-1}v|_{\bar{g}} \leq 1$. Suppose that on C it holds that

$$\int_{t_0}^{t} |\nabla N|_{\bar{g}} dt = c \tag{2.25}$$

and

$$K_{ij} v^i v^j \geq k > 0; \tag{2.26}$$

then

$$z(t) \leq z(t_0) \frac{N_M}{N(t_0)} (\exp c) \exp[-k(t - t_0)]; \tag{2.27}$$

therefore z is integrable on C.

The proof of the corollary rests on the fact that

$$|P_{ij} v^i v^j| \leq |P|_{\bar{g}} |v|^2_{\bar{g}} \leq N^2 |P|_{\bar{g}}. \tag{2.28}$$

□

Remark 2.5 *Weaker sufficient conditions for incompleteness in the same spirit can be formulated, but with a less clear geometric meaning.*

3 Congruence of timelike curves

Before giving a succinct account of the Hawking–Penrose singularity theorems, we review some properties of a congruence of curves.

3.1 *Definitions*

Let u be a timelike [respectively a null, a causal] vector field on the $n+1$-dimensional spacetime (V, g). Its trajectories are called a timelike [respectively a null, a causal] congruence. They are defined by integrating the differential system

$$\frac{dx^\alpha}{d\lambda} = u^\alpha(x). \tag{3.1}$$

It is easy to see by a change of the parameter λ on the integral curves that the vector fields u and au, with a a smooth function which has no zero on V, define the same congruence.

3.1.1 *Spacetime properties*

Definition 3.1 *The **rotational** tensor of a congruence with cotangent vector u is the antisymmetric tensor $\Omega = du$ with components*

$$\Omega_{\alpha\beta} := \nabla_\alpha u_\beta - \nabla_\beta u_\alpha. \tag{3.2}$$

We remark that the definition of Ω depends not only on the trajectories of u but on u itself.

If u is the gradient of a scalar function f, $u_\alpha = \partial_\alpha f$, then $\Omega_{\alpha\beta} = 0$. The converse is locally true, by elementary analysis.

Definition 3.2 *The congruence is called **irrotational** if the rotational tensor of a tangent vector field to its trajectories is zero.*

Theorem 3.3 *A congruence in an $n+1$-dimensional spacetime which is orthogonal to a family of smooth n-submanifolds $f =$ constant is irrotational.*

Proof We can take as tangent vector field to the trajectories of the congruence the gradient of f. □

Theorem 3.4 *If a timelike geodesic congruence is irrotational (i.e. $\Omega = 0$ when the curves are parametrized by arc length or canonical parameter) on some transversal hypersurface it remains irrotational.*

Proof The Ricci identity gives

$$u^\lambda \nabla_\lambda \nabla_\alpha u_\beta \equiv u^\lambda \nabla_\alpha \nabla_\lambda u_\beta + u^\lambda R_{\lambda\alpha\beta\gamma} u^\gamma. \tag{3.3}$$

If the flow of u is geodesic and canonically parametrized (that is, if $u^\lambda \nabla_\lambda u_\beta = 0$), the above identity implies that:

$$u^\lambda \nabla_\lambda \nabla_\alpha u_\beta = -\nabla_\alpha u^\lambda \nabla_\lambda u_\beta + u^\lambda R_{\lambda\alpha\beta\gamma} u^\gamma. \tag{3.4}$$

By the definition of Ω we have

$$u^\lambda \nabla_\lambda \Omega_{\alpha\beta} = u^\lambda \nabla_\lambda \{\nabla_\alpha u_\beta - \nabla_\beta u_\alpha\};$$

hence, using the symmetries of the Riemann tensor,

$$\begin{aligned} u^\lambda \nabla_\lambda \Omega_{\alpha\beta} &= -\nabla_\alpha u^\lambda \nabla_\lambda u_\beta + \nabla_\beta u^\lambda \nabla_\lambda u_\alpha \\ &= (\nabla_\lambda u_\alpha - \nabla_\alpha u_\lambda)\nabla_\beta u^\lambda + \nabla_\alpha u^\lambda (\nabla_\beta u_\lambda - \nabla_\lambda u_\beta). \end{aligned}$$

We see that $\Omega_{\alpha\beta}$ satisfies the linear and homogeneous differential equation along the geodesic

$$u^\lambda \nabla_\lambda \Omega_{\alpha\beta} = \nabla_\alpha u^\lambda \Omega_{\beta\lambda} + \nabla_\beta u^\lambda \Omega_{\lambda\alpha}, \tag{3.5}$$

from which the conclusion follows. □

A particular case where the theorem applies is when the geodesics originate from one point. We have met this case in Section XII.7.

3.1.2 *Expansion and space tensors*

The **expansion** θ of a congruence with tangent vector u is the divergence

$$\theta := \nabla_\alpha u^\alpha.$$

Let u be a field of unit timelike vectors. Its covariant derivative is orthogonal to it:

$$u^\beta \nabla_\alpha u_\beta = 0, \quad \text{since} \quad u^\beta u_\beta = -1. \tag{3.6}$$

One defines the velocity of rotation and the shear of the congruence by decomposing the covariant derivative ∇u into tensors orthogonal to u, all but one being traceless.

Recall that the positive definite symmetric projection tensor π associated with a unit vector u is

$$\pi_{\alpha\beta} := g_{\alpha\beta} + u_\alpha u_\beta, \quad \text{hence} \quad \pi^\alpha_\alpha = n. \tag{3.7}$$

Lemma 3.5 *The tensor field ∇u, with u a unit timelike vector field admits the decomposition*

$$\nabla_\alpha u_\beta \equiv \omega_{\alpha\beta} + \sigma_{\alpha\beta} + \frac{1}{n}\theta \pi_{\alpha\beta} - (u^\lambda \nabla_\lambda u^\alpha) u_\beta \tag{3.8}$$

where ω is the antisymmetric tensor projection of the rotational tensor of u on the tangent space orthogonal to u, called the velocity of rotation, or ***vorticity*** *tensor, given by*

$$\omega_{\alpha\beta} := \frac{1}{2}\pi^\rho_\alpha \pi^\sigma_\beta \{\nabla_\rho u_\sigma - \nabla_\rho u_\sigma\}. \tag{3.9}$$

The symmetric tensor σ, orthogonal to u, is called the ***shear****, which has a zero trace,*

$$\sigma_{\alpha\beta} := \frac{1}{2}\pi^\rho_\alpha \pi^\sigma_\beta \{\nabla_\rho u_\sigma + \nabla_\rho u_\sigma\} - \frac{1}{n}\theta \pi_{\alpha\beta}, \qquad \sigma^\alpha_\alpha = 0,$$

Exercise. Prove this identity.

For a general congruence the vorticity and shear can be written, using the expression of the projection tensor π,

$$\omega_{\alpha\beta} := \frac{1}{2}\{\nabla_\alpha u_\beta - \nabla_\beta u_\alpha + (u^\lambda \nabla_\lambda u_\alpha)u_\beta - (u^\lambda \nabla_\lambda u_\beta)u_\alpha\}.$$

and,

$$\sigma_{\alpha\beta} := \frac{1}{2}\{\nabla_\alpha u_\beta + \nabla_\beta u_\alpha + (u^\lambda \nabla_\lambda u^\alpha)u_\beta + (u^\lambda \nabla_\lambda u^\beta)u_\alpha\} - \frac{1}{n}\theta\pi_{\alpha\beta}, \quad \sigma^\alpha_\alpha = 0.$$

When the congruence is a **geodesic congruence**, i.e. when

$$u^\lambda \nabla_\lambda u^\alpha = 0$$

the tensors ω and σ simplify respectively to the rotational of u and a modified conformal Lie derivative of g in the direction of u,

$$\omega_{\alpha\beta} = \frac{1}{2}\{\nabla_\alpha u_\beta - \nabla_\beta u_\alpha\},$$
$$\sigma_{\alpha\beta} = \frac{1}{2}\{\nabla_\alpha u_\beta + \nabla_\beta u_\alpha\} - \frac{1}{n}\theta\pi_{\alpha\beta}.$$

3.2 *Geodesic deviation*

Let C_σ be a 1-parameter **family of timelike geodesics parametrized** by their arc length s, $C_\sigma(s) = \psi(\sigma, s)$. Denote by $u = \frac{\partial\psi}{\partial s}$ the tangent vector to C_σ and by $h = \frac{\partial\psi}{\partial\sigma}$ the vector which characterizes the infinitesimal displacement of C_σ for a given s. These two vectors commute, as partial derivatives of a function ψ, i.e.

$$u^\alpha \nabla_\alpha h^\beta - h^\alpha \nabla_\alpha u^\beta = 0. \tag{3.10}$$

Differentiating this relation in the direction of u, and using the parallel transport of u, gives

$$u^\lambda u^\alpha \nabla_\lambda \nabla_\alpha h^\beta - u^\lambda \nabla_\lambda h^\alpha \nabla_\alpha u^\beta - h^\alpha u^\lambda \nabla_\lambda \nabla_\alpha u^\beta = 0; \tag{3.11}$$

that is, by the Ricci formula

$$u^\lambda u^\alpha \nabla_\alpha h^\beta - u^\lambda \nabla_\lambda h^\alpha \nabla_\alpha u^\beta - h^\alpha u^\lambda (\nabla_\alpha \nabla_\lambda u^\beta + R_{\lambda\alpha}{}^\beta{}_\mu u^\mu) = 0. \tag{3.12}$$

It holds that, since u is parallely transported and commutes with h

$$h^\alpha u^\lambda \nabla_\alpha \nabla_\lambda u^\beta = -h^\alpha \nabla_\alpha u^\lambda \nabla_\lambda u^\beta = -u^\alpha \nabla_\alpha h^\lambda \nabla_\lambda u^\beta, \tag{3.13}$$

therefore the relation (3.11) simplifies to

$$u^\lambda u^\alpha \nabla_\lambda \nabla_\alpha h^\beta = h^\alpha u^\lambda u^\mu R_{\lambda\alpha}{}^\beta{}_\mu, \tag{3.14}$$

which can also be written

$$\nabla^2_{u^2} h^\beta := \frac{D^2}{Ds^2} h^\beta = h^\alpha u^\lambda u^\mu R_{\lambda\alpha}{}^\beta{}_\mu. \tag{3.15}$$

This equation, linking with the curvature the rate of acceleration of distance between nearby geodesics, is called the equation of **geodesic deviation**. It is a fundamental equation for interpreting the effects of gravitation in the Einstein theory.

3.3 *Raychauduri equation*

The derivative of the expansion $\theta := \nabla_\alpha u^\alpha$ of a geodesic flow in the direction of the vector u is obtained by taking the trace of Equation (3.4). We find

$$\frac{d\theta}{ds} := u^\lambda \nabla_\lambda \theta = -\nabla_\alpha u^\lambda \nabla_\lambda u^\alpha - R_{\lambda\gamma} u^\lambda u^\gamma.$$

Hence, using the decomposition (3.8) of ∇u when u is a geodesic flow, together with the symmetry of σ and antisymmetry of ω,

$$\frac{d\theta}{ds} = -\frac{\theta^2}{n} - \sigma_{\alpha\beta}\sigma^{\alpha\beta} + \omega_{\alpha\beta}\omega^{\alpha\beta} - R_{\lambda\gamma} u^\lambda u^\gamma.$$

This equation is known as the **Raychauduri equation**.

The following theorem is an easy consequence of the Raychauduri equation; it is a symptom of the attractive character of gravitation.

Theorem 3.6 *Let u be the tangent vector field to an irrotational timelike geodesic congruence. Suppose that $R_{\alpha\beta}u^\alpha u^\beta \geq 0$ (i.e. the spacetime satisfies the **strong energy condition**) and the expansion θ takes a negative value θ_0 at some point on a geodesic of the congruence. Then θ goes to $-\infty$ along this geodesic within the proper time $-n(\theta_0)^{-1}$.*

Proof Under the given hypotheses the Raychauduri equation implies

$$\frac{d\theta}{ds} \leq -\frac{1}{n}\theta^2, \tag{3.16}$$

because $\omega = 0$ and σ is orthogonal to u, and hence has non-negative g-pseudo-norm. Therefore θ will be decreasing along the geodesic, and hence remains negative if it starts negative. Set $\tilde{\theta} = -\theta$. We deduce from the above inequality

$$\frac{d\tilde{\theta}}{ds} \geq \frac{1}{n}\tilde{\theta}^2, \quad \text{i.e.} \quad \frac{ds}{d\tilde{\theta}} \leq \frac{n}{\tilde{\theta}^2}. \tag{3.17}$$

The value $\tilde{\theta} = +\infty$ will be attained for a finite value s_∞ of the proper time s, because (3.17) implies that, for all $\tilde{\theta} \geq \tilde{\theta}_0$,

$$s(\tilde{\theta}) \leq s(\tilde{\theta}_0) + \int_{\tilde{\theta}_0}^{\tilde{\theta}} \frac{n}{\tilde{\theta}^2} d\tilde{\theta} \leq s_\infty = s_0 + \frac{n}{\tilde{\theta}_0}, \qquad \tilde{\theta}_0 = -\theta_0.$$

□

Remark 3.7 *The fact that the absolute value of θ becomes infinite does not in itself signal a singularity of the space time. It might be only a singularity of the congruence.*

3.4 *Null geodesic congruence*

A null geodesic can be canonically parametrized, but the canonical parameter λ is defined only up to an affine transformation[5] $\lambda \to \lambda_0 + c\lambda$ because there is no intrinsic way to normalize a null vector ℓ, and thereby no intrinsic way of adjusting the scale of λ on different null geodesics.

The vector space P_x orthogonal to a null vector $\ell \in T_xV$ is n-dimensional, as when the vector is not null, but it now contains the vector ℓ itself; indeed P_x is the tangent space to the null cone at x along ℓ. There is no canonical way to decompose a vector, or tensor, into components parallel and orthogonal to ℓ. There are several ways to establish equations of geodesic deviations for null geodesics (see Hawking and Ellis, 1973; Wald, 1984), and an equation analogous to the Raychauduri equation. When ℓ is a null vector the condition $R_{\alpha\beta}\ell^\alpha\ell^\beta \geq 0$ is equivalent to $T_{\alpha\beta}\ell^\alpha\ell^\beta \geq 0$.

4 First singularity theorem

The study of conjugate points is fundamental for the singularity theorems, proven in the early 1970s by Hawking and Penrose.

4.1 *Conjugate points*

Definition 4.1

1. *A Jacobi field on a geodesic C is a vector field h defined along C, and orthogonal*[6] *to C, which satisfies the second-order linear differential equation (3.15) of geodesic deviation.*

2. *Two points a and b on a geodesic C are said to be conjugate if there is a Jacobi field along C which vanishes at a and b.*

Theorem 4.2

1. *The point b is conjugate to a along the geodesic C if and only if the expansion θ of the geodesic congruence starting from a tends to $-\infty$ at b along C.*

2. *A smooth timelike curve C connecting a to b realizes a local maximum of the length ℓ between these points if and only if it is a geodesic with no point between a and b which is conjugate to a.*

Proof The deviation vectors h along C are orthogonal to the tangent u to C. We choose an orthonormal frame along C with time vector u. In such a frame

[5] This is why λ is called an affine parameter.

[6] If C is not a null geodesic. If C is a null geodesic one must choose a canonically defined vector transversal to C to define the geodesic deviation.

we have $h^0 = h_0 = 0$. For further simplification, we choose a frame which is parallely transported along C so that h satisfies along C the following equation

$$u^\lambda \nabla_\lambda h^i = u^\lambda \partial_\lambda h^i = \frac{dh^i}{ds}$$

We also choose this basis $h_{(j)}$, $j = 1, \ldots, n$, to satisfy the initial conditions

$$\left\{ \frac{d}{ds} h^i_{(j)} \right\} (0) = \delta^i_j.$$

By the definition of a deviation vector we have

$$\frac{Dh^i_{(j)}}{Ds} \equiv u^\lambda \nabla_\lambda h^i_{(j)} = h^k_{(j)} \nabla_k u^i$$

Therefore in the chosen frame, denoting by A the $n \times n$ matrix with elements $h^i_{(j)}$ and by B the matrix with elements $\nabla_k u^i$

$$\frac{dA}{ds} = AB, \quad \text{hence} \quad A^{-1} \frac{dA}{ds} = B$$

This equality implies, by classical algebra

$$\frac{1}{\det A} \frac{d(\det A)}{ds} = \frac{d(\log \det A)}{ds} = traceB$$

If there exists a Jacobi field along C vanishing at a and b, where $s = s_b$, then $\det A$ tends to zero as s tends to s_b on C. If the conjugate point is isolated $\frac{d}{ds}(\det A)$ remains away from zero; then $traceB$ tends to $-\infty$.

We complete the proof by remarking that in an orthonormal frame with time axis u it holds that

$$\theta := \nabla_\alpha u^\alpha = \omega^\alpha_{\alpha 0} = \omega^i_{i0} = \nabla_i u^i := trB.$$

2. Suppose C is a timelike geodesic without conjugate points between a and b. The matrix A defined above in the parallely transported frame is then non-singular along C. We define a vector field w along C by giving its components in this frame, namely we can set

$$h^i = A^i_j w^j$$

where w is a vector orthogonal to u which vanishes at a and b. One can then prove (cf. Hawking and Ellis, proposition 4 8 5 extended to arbitrary n) that the second derivative of the length functional satisfies the inequality

$$\ell''(h, h) = -\int_{s_a}^{s_b} \Sigma_{i=1}^n \left(A_{ij} \frac{dw^j}{ds} \right)^2 ds < 0$$

for all h non-identically zero. The result follows. □

4.2 *Incompleteness theorem*

One can now prove the following theorem, which involves only conditions to be satisfied by the sources and the initial data.

Theorem 4.3 *Let (V, g) be a globally hyperbolic spacetime such that $R_{\alpha\beta}v^\alpha v^\beta \geq 0$ for all timelike vectors v, and assume there exists a C^2 Cauchy surface S with mean curvature $\tau := trK \geq c > 0$, c some given positive number. Then there exists no future-directed timelike curve from S with length greater than nc^{-1}.*

Proof By definition the mean extrinsic curvature K of S at a point a is $K = -\theta$, with θ the expansion at that point of the (future-directed) geodesic congruence normal to S. By the theorem of the previous section this expansion becomes $-\infty$ along a geodesic of the congruence within a proper time no greater than nc^{-1}, inflicting thus a conjugate point b to a along this geodesic, preventing a geodesic segment from a to b from realizing a local maximum of length as soon as this length is greater than nc^{-1}. But, since S is a Cauchy surface, its future is globally hyperbolic. In a globally hyperbolic spacetime if b is in the strict future of a there exists a timelike geodesic between a and b realizing the local maximum of length. This length cannot be greater than nc^{-1}. □

Remark 4.4 *An analogous theorem is obviously obtained by exchanging future and past and switching the sign of c in the hypothesis. Such a theorem states the existence of a past singularity at a distance at most $n|c|^{-1}$.*

5 Trapped surfaces and singularities

5.1 *Trapped surfaces*

The Eddington–Finkelstein spacetime has a curvature singularity when r tends to zero. This singularity is announced by the fact that the 2-spheres $t = constant$, $r = constant < 2m$ (in Schwarzschild coordinates) are **trapped surfaces**. To define, generically, a trapped surface one uses the definition of the expansion of a congruence of null geodesics given by the trace of the null second fundamental form.

We recall[7] the definition of the second fundamental form of a submanifold Σ of a pseudo- Riemannian manifold (V, g). Let u and v be tangent vector fields to Σ. The covariant derivative in the spacetime metric g of v in the direction of u is a well defined vector $\nabla_u v$ in the tangent space TV. One shows that the normal component $(\nabla_u v)^{\$}_x$ of $\nabla_u v$ at x depends only on the values u_x and v_x. It defines a quadratic form on $T_x\Sigma$ with value in the $g-$orthogonal space $(T_x\Sigma)^{\$}$ to $T_x\Sigma$, supposed to be non-isotropic. It is called the second fundamental form of Σ.

[7] See for instance CB-DM1 VB 3.

Suppose for example that Σ is a hypersurface M of V, i.e. an n-dimensional submanifold of an $n+1$-dimensional spacetime. One sets

$$(\nabla_u v)^{\$} = n_\alpha (\nabla_u v)^\alpha, \tag{5.1}$$

with n the future-directed unit normal to Σ. We consider a Cauchy adapted frame corresponding to M and V. The components of u and v, tangent to M, are then such that $u^0 = v^0 = 0$. Therefore the components of $\nabla_u v$ are

$$(\nabla_u v)^j = u^i \nabla_i v^j = u^i (\partial_i v^j + \Gamma^j_{ik} v^k) \equiv u^i \bar{\nabla}_i v^j, \tag{5.2}$$

vector tangent to M which depends only on u, v on M and the induced metric; and the component normal to M, given by the quadratic form in u and v

$$K(u,v) := n_\alpha (\nabla_u v)^\alpha = -N(\nabla_u v)^0 = -N u^i \nabla_i v^0 = -N u^i \omega^0_{ik} v^k. \tag{5.3}$$

The covariant symmetric 2-tensor K defined by the quadratic form $K(u,v)$ coincides with the extrinsic curvature defined in Section 6.3.

Suppose now that (V,g) is a spacetime of dimension $n+1$, and that Σ is of dimension $n-1$ and spacelike; then $(T_x\Sigma)^{\$}$ has dimension 2 and is timelike. Such a 2-plane cuts the future light cone at x along two future null directions. We choose two future null vectors denoted ℓ^+ and ℓ^-; supported by these directions we define the two null second fundamental forms of Σ by

$$K^+(u,v) := g(\ell^+, \nabla_u v) \equiv \ell^+_\alpha (\nabla_u v)^\alpha, \quad K^-(u,v) := \ell^-_\alpha (\nabla_u v)^\alpha. \tag{5.4}$$

These forms are defined up to scaling by a positive factor of the vectors ℓ^+ and ℓ^- on the future null directions. Note that, since ℓ_+ and ℓ_- are future directed, the g scalar product (ℓ_+, ℓ_-) is negative.

Denote by e_a, $a = 2, \ldots, n$ a reference frame on Σ, hence orthogonal to ℓ_+ and ℓ_-. The component $K^+_{ab} := K(e_a, e_b)$ is obtained by the scalar product of ℓ^+ with $\nabla_{e_a} e_b$ We use the orthogonality of ℓ^+ with e_b to find

$$K^{\dot{+}}_{ab} = g(\nabla_{e_a} e_b, \ell^+) = g(e_b, \nabla_{e_a} \ell^+) \equiv \nabla_a \ell^+_b, \tag{5.5}$$

and a similar formula for K^-_{ab}.

We denote by χ^+ and χ^-, the **null mean curvatures** of Σ, given by the traces of K^+ and K^- in the metric $h = (h_{ab})$ induced by g on Σ,

$$\chi^+ = h^{ab} \nabla_a \ell^+_b, \quad \chi^- = h^{ab} \nabla_a \ell^-_b. \tag{5.6}$$

We suppose that Σ is an oriented $n-1$ submanifold of an n-dimensional spacelike submanifold M of V, so we can continuously orient the normal ν to Σ in M, and distinguish the interior and exterior of Σ in M.

Lemma 5.1 *The null mean curvatures are the expansions of the null geodesic congruences defined by ℓ^+ and ℓ^-.*

Proof Let n and ν be respectively the future unit normal to M, and the unit normal to Σ tangent to M and pointing outwards. Consider the null vectors ℓ^+ and ℓ^- given by

$$\ell^+ = n + \nu, \quad \ell^- = n - \nu, \quad \text{then} \quad \ell^+_\alpha \ell^\alpha_- = -2. \tag{5.7}$$

Denote by h the spacetime tensor

$$h := g + n \otimes n - v \otimes \nu \equiv g + \frac{1}{2}(\ell^+ \otimes \ell^- + \ell^- \otimes \ell^+). \tag{5.8}$$

That is,

$$h^{\alpha\beta} = g^{\alpha\beta} + n^\alpha n^\beta - \nu^\alpha \nu^\beta = g^{\alpha\beta} - \frac{1}{2}(\ell^\alpha_+ \ell^\beta_- + \ell^\alpha_- \ell^\beta_+). \tag{5.9}$$

Take as a V-reference frame along Σ the $n+1$ vectors (ℓ_+, ℓ_-, e_a). The components h^{ab}, $a, b = 2, \ldots, n$ of this tensor coincide with the components of the induced metric on Σ so denoted before, and it is easy to check that its other components vanish in the considered frame. Therefore

$$\chi^+ = h^{\alpha\beta} \nabla_\alpha \ell^+_\beta. \tag{5.10}$$

But it holds that

$$\ell^\alpha_+ \nabla_\beta \ell^+_\alpha = 0, \tag{5.11}$$

since ℓ^+ is a null vector, and

$$\ell^\alpha_+ \nabla_\alpha \ell^+_\beta = 0, \tag{5.12}$$

if ℓ^+ is tangent to a geodesic congruence. Under these hypotheses we have therefore

$$\chi^+ = g^{\alpha\beta} \nabla_\alpha \ell^+_\beta \equiv \nabla_\alpha \ell^\alpha_+. \tag{5.13}$$

The same proof holds for ℓ^-. □

We can now give the definition of a trapped surface, which implies that the areas of the inward and outward shells of light emitted by a trapped surface are both decreasing.

Definition 5.2 *A trapped surface in a spacetime (V, g) is a compact $n-1$-dimensional spacelike submanifold Σ such that both the null mean curvatures χ^+ and χ^- are negative*[8] *on Σ, i.e. the null geodesic congruences both converge.*

We derive an expression for χ^+ which is useful in the formulation of boundary value problems.

[8] To agree with most recent papers we have taken a signature convention opposite to the one taken for the mean curvature of a hypersurface, by choosing $\ell^\pm$ to be future directed.

Lemma 5.3 *The mean null curvature χ^+ relative to ℓ^+ is given by*

$$\chi^+ = -tr_{\bar{g}}K + K(\nu,\nu) + tr_h k, \tag{5.14}$$

where $tr_{\bar{g}}K$ is the mean extrinsic curvature of M in (V,g) and $tr_h k$ is the mean extrinsic curvature of Σ in $(M,\bar{g})$.

Proof It holds that

$$g^{\alpha\beta}\nabla_\alpha \ell^+_\beta = (g^{\alpha\beta} + n^\alpha n^\beta - \nu^\alpha\nu^\beta)\nabla_\alpha(n_\beta + \nu_\beta). \tag{5.15}$$

The tensor $g^{\alpha\beta} + n^\alpha n^\beta$ is the contravariant form of the metric $\bar{g}$ induced by g on M, hence, since here n is future oriented, we have

$$(g^{\alpha\beta} + n^\alpha n^\beta)\nabla_\alpha n_\beta = -tr_{\bar{g}}K. \tag{5.16}$$

Since ν is tangent to M, i.e. $\nu^0 = 0$ in a Cauchy adapted frame,

$$\nu^\alpha\nu^\beta\nabla_\alpha n_\beta = -\nu^i\nu^j K_{ij}. \tag{5.17}$$

Finally

$$(g^{\alpha\beta} + n^\alpha n^\beta - \nu^\alpha\nu^\beta)\nabla_\alpha\nu_\beta = h^{ab}\bar{\nabla}_a\nu_b = tr_h k \tag{5.18}$$

because $\nabla_a\nu_b = \bar{\nabla}_a\nu_b$, where $\bar{\nabla}$ is the covariant derivative in the metric $\bar{g}$.

The obtained formulas give the result. □

The following definition is particularly useful in the study of black holes.

Definition 5.4 *The manifold Σ is called an **apparent horizon** if on Σ*

$$\chi^+(\Sigma) := -tr_{\bar{g}}K + K(\nu,\nu) + tr_h k = 0 \tag{5.19}$$

A sliced spacetime is called **time symmetric** if it admits a slice M_0 with zero extrinsic curvature. The solution of the Cauchy problem is then invariant by time reversal.

At a moment of time symmetry, $K = 0$, the expansion χ^+ reduces to $tr_h k$, and we have therefore the following lemma.

Lemma 5.5 *An apparent horizon in a space manifold M which is a moment of time symmetry is a submanifold with zero mean extrinsic curvature.*

Remark 5.6 *An $n-1$-submanifold with zero mean extrinsic curvature in a Riemannian n-manifold is a minimal submanifold, giving a local minimum of the volume of $n-1$-submanifolds.*

Exercise. Prove that the 2-spheres Σ given by $x^0 := t = T$, $x^1 := r = R > 2m$ in the Schwarzschild spacetime with metric

$$ds^2 = -(1-2mr^{-1})dt^2 + (1-2mr^{-1})^{-1}dr^2 + r^2(sin^2\theta d\phi^2 + d\theta^2).$$

are non-trapped.

Prove that in the black hole extension of the Schwarzschild spacetime, again in Schwarzschild coordinates, the same 2-manifolds but with $R < 2m$ are trapped surfaces.

5.2 *Singularities linked to trapped surfaces*

We give only two of the singularity theorems proved by Penrose and Hawking. More theorems and proofs can be found in the books by Hawking and Ellis (1973) and Wald (1984).

The first theorem, due to Penrose (1965)[9], concerns data on non-compact manifolds.

Theorem 5.7 *If the initial data set $(M, \bar{g}, K)$, with M non-compact, possesses a trapped surface, then its maximal future globally hyperbolic development is incomplete.*

In fact, it can be proved that at least one of the future-directed null geodesics orthogonal to the trapped surface has a bounded affine length. The theorem holds for Einstein equations with sources such that $R_{\alpha\beta}k^\alpha k^\beta$ for all null vectors, hence for sources satisfying the weak or the strong energy condition. Work by Christodoulou (2008) gives the conditions for a trapped surface to form under the evolution of smooth data possessing no such surface.

The second theorem applies to spacetimes which Penrose calls "generic", with Riemann tensor such that each causal geodesic possesses at least one point at which $X^\gamma X^\delta X_{[\alpha}R_{\beta]\gamma\delta[\lambda}X_{\mu]} \neq 0$ for every causal vector X ([...] denotes antisymmetrization).

Theorem 5.8 *Let (V, g) be a generic spacetime with no closed timelike curve. Suppose that for every causal vector X, $R_{\alpha\beta}X^\alpha X^\beta \geq 0$. Then (V, g) is causal geodesic incomplete if it possesses a trapped surface.*

See the proof in the quoted books. Remark that the second theorem involves a hypothesis on the spacetime itself, not only on initial data.

6 Black holes

We studied in Chapter 4 the Schwarzschild spacetime and black hole, which remain a fundamental conceptual guide for the understanding of Einsteinian spacetimes. Black holes are one of the most striking features of our Universe, predicted by the Einstein equations. Since the observation by astronomers of the first candidate black hole in the 1960s many more black holes have revealed themselves by the motion of surrounding stars or an exceptionally bright accretion ring which reveals the presence of an enormous mass confined in a region too small for any known matter.

6.1 *Definitions*

An intuitive physical definition of a black hole is easy to give: a black hole is, roughly speaking, a region of spacetime in which the gravitational field is so

[9] Penrose, R. (1965) *Phys. Rev. Lett.*, **14**, 57.

strong that light cannot get out. However, a precise mathematical definition is more difficult and still controversial. Light does not escape from the future of a point, but such regions cannot be considered as black holes. A definition of black holes common at present in mathematical Relativity uses conformal completion and is given in Appendix VI. This definition has the inconvenience of requiring global properties of spacetime, and it does not apply when the space manifolds are compact.

A possible similar definition, which does not appeal to conformal compactification is the following.

Definition 6.1 *The **black hole region** $\mathcal{B}$ of a spacetime (V, g) is the complement of the past of the set covered by the null geodesics which have an infinite future canonical parameter.*

*The **event horizon** H is the boundary $\partial\mathcal{B}$ of the black hole region $\mathcal{B}$ in the spacetime V.*

We have seen in the Finkelstein–Schwarzschild example of a black hole that the limit $r = 2m$ of the trapped surfaces, $r = r_0$ with $r_0 < 2m$, is an **apparent horizon** which generates, under time evolution, the **event horizon**, that is the boundary of the black hole region $r < 2m$, from within which no signal can escape.

Geometrical and physical intuition lead us to think that these properties hold in general, under appropriate reasonable assumptions: namely the boundary H of the black hole region is null, generated by null geodesics, and a section of H by a space manifold is an apparent horizon. It is difficult to reduce these "reasonable assumptions" to a minimum. We will not attempt it. We will simply try to give the hypotheses which are made when we state a theorem involving black holes.

We extend the definition of asymptotically strongly predictable spacetimes to our definition of black holes, adding the word "regularly" to include a property of the event horizon which can be proved to hold under some reasonable hypotheses[10].

Definition 6.2 *A black hole spacetime is called **asymptotically strongly predictable** if the complement of the closure of the black hole region is globally hyperbolic. It is called **regularly asymptotically strongly predictable** if, in addition, the event horizon is generated by null geodesics.*

The singularity theorem 5.7 implies that a black hole spacetime $\mathcal{B}$ cannot be embedded in a Lorentzian manifold V which is globally hyperbolic and complete V if $V - \mathcal{B}$ contains a trapped surface, $\chi^+ < 0$.

One can define similarly the white hole region of a spacetime by exchanging future and past and black by white.

[10] We proved in Chapter 12 that in a globally hyperbolic spacetime the boundary of the past, or future, of a point is generated by null geodesics.

There are many theorems and conjectures about black holes, difficult to prove and even often to state in a precise way. We state the Hawking area theorem, basic for the thermodynamic interpretation of the properties of black holes and their link with quantum phenomena, and the Penrose inequality proved in the time symmetric case by Huisken and Ilmanen and in a more general case by Bray through sophisticated mathematics.

6.2 *The Hawking area theorem*

Theorem 6.3 *Let (V, g) be an asymptotically flat, regularly asymptotically strongly predictable Einsteinian spacetime with sources such that $R_{\alpha\beta}\ell^\alpha\ell^\beta \geq 0$ for all null vectors ℓ. Let M_1 and M_2 be two Cauchy surfaces in the globally hyperbolic region, with M_2 in the future of M_1, i.e. $M_2 \subset I^+(M_1)$. Denote by $\mathcal{H}_1$ and $\mathcal{H}_2$ the intersections of the event horizon H with M_1 and M_2. Then the area of $\mathcal{H}_2$ is greater than the area of $\mathcal{H}_1$.*

Proof See Wald, Chapter 8, and recent advances with relaxed hypotheses and precise conclusions in Chruściel, Delay, Galloway and Howard. All these authors use definitions through an unphysical conformal spacetime, although this unphysical formulation, which restricts the study to asymptotically Minkowskian spacetime, is probably not necessary. □

6.3 *The Riemannian Penrose inequality. Case $n = 3$*

The Riemannian Penrose inequality is the time symmetric case of a more general conjecture of Penrose for the mass of a black hole and the area of its boundary.

The Riemannian Penrose inequality has the advantage of admitting a formulation in purely geometrical terms, because in the time symmetric case the boundary of a black hole is a minimal 2-surface in the three-dimensional space manifold. It was proved at the beginning of this century, by Huisken and Ilmanen and also by Bray by a different method[11].

To understand the statement of the theorem one must know the following definition.

An asymptotically Euclidean Riemannian manifold with possibly non-empty interior and an end diffeomorphic to $R^3 - K$, K a compact set, is called **harmonically flat at infinity** if the metric takes in the end the form

$$g_{ij} = u^4(x)\delta_{ij}, \tag{6.1}$$

where u satisfies the flat space Laplace operator in the coordinates of the end, and goes to a constant at infinity. The classical expansion of u in spherical

[11] See e.g. Bray, H. and Chruściel, P. (2004) The Penrose inequality. In *The Einstein Equations and the Large Scale Behaviour of Spacetime* (eds. P. Chruściel and H. Friedrich), Birkhäuser, and references therein to the original articles and previous heuristic proofs.

harmonics shows that u admits then the following asymptotics as $|x|$ tends to infinity

$$u(x) = a + \frac{b}{|x|} + O\left(\frac{1}{|x|^2}\right). \tag{6.2}$$

Schoen and Yau showed when proving the positive mass theorem that, given any $\varepsilon > 0$, it is always possible to perturb an asymptotically Euclidean manifold to become harmonically flat at infinity with the metric changing by less than ε pointwise, keeping a non-negative scalar curvature and with the ADM mass changing by less than ε (even stronger statements are obtained by Corvino (see Section VII.13).

Theorem 6.4 *Let (M, g) be a complete smooth three-dimensional manifold with non-negative scalar curvature which is harmonically flat at infinity with ADM mass m and which has an outermost*[12] *minimal surface Σ_0 with area A_0. Then*

$$m \geq \sqrt{\frac{A_0}{16\pi}} \tag{6.3}$$

with equality if and only if outside of its outermost minimal surface (M, g) is isometric to the Schwarzschild spacetime $(R^3 \cap \{|x| > m\}, u = (1 + \frac{m}{2|x|})^4)$.

The proof uses among other arguments the inverse mean curvature flow.

7 Weak cosmic censorship conjectures

7.1 *Naked singularity*

The original idea of Penrose[13], coming from the study of spherical gravitational collapse where a black hole forms, hiding the singularity from timelike observers, is the conjecture that generic Einsteinian spacetimes with physically reasonable sources do not admit any **naked singularity**; that is, a singularity visible by an observer. With the definition of singularity by incompleteness, trivial counter-examples are obtained by cutting out regions of any spacetime, for example Minkowski spacetime. Therefore the mathematics must be more precise to grasp the physical content. In the previous chapter we considered complete initial data, and stated the strong cosmic principle for vacuum Einstein space times. We give the following definition of a nakedly singular spacetime.

Definition 7.1 *An inextendible spacetime (V, g) is said to be future nakedly singular if it admits a future inextendible causal curve which lies entirely in the past of some point $x \in V$.*

[12] It is a minimal surface boundary of the union of the open regions of M bounded by all the minimal surfaces in (M, g).

[13] Penrose, R. (1979) In *An Einstein Centenary Survey* (eds. S. Hawking and W. Israel), Cambridge University Press.

We formulate the conjecture as follows

No naked singularity conjecture. *An inextendible, generic*[14]*, Einsteinian spacetime with physically reasonable sources admits no naked singularity.*

Remark that the Big Bang is not a counter-example to this conjecture; it has no past, and hence does not correspond to any future inextendible causal curve.

A generic spacetime can be understood as a spacetime that is stable, in some sense to be defined, under small perturbations. Reasonable sources are physical sources which have a hyperbolic, causal evolution and do not have their own singularities (shocks, shell crossings, etc.).

It results from the definition of global hyperbolicity that a nakedly singular spacetime is not globally hyperbolic. Penrose proves that, conversely if a space is strongly causal and is not nakedly singular it is globally hyperbolic. This led to the strong cosmic censorship conjecture (see Chapter 12), saying in a rather loose way that all "reasonable" Einsteinian spacetimes ought to be globally hyperbolic, in particular they should never be nakedly singular. For the same reason as the Big Bang, the Schwarzschild metric with $m < 0$ on the manifold $(R^3 - \{0\}) \times R$ is not a counter example to the "non-naked singularity" conjecture. Remark also that the Schwarzschild metric with $m < 0$ is not considered as a meaningful physical metric, because no reasonable physical situation is modelled by such a metric on a manifold $(R^3 - \omega) \times R$, with ω a neighbourhood of 0, since the positive energy theorem (see Related paper) excludes a metric with such a behaviour at spacelike infinity for sources with non-negative energy.

7.2 *Weak cosmic censorship*

The weak cosmic censorship conjecture says, in loose terms, that if a singularity exists, it cannot be seen by "asymptotic observers". That is, there is no light ray which "escapes to infinity" from points where the singularity can be observed. A usual formulation is through the definition of future null infinity $\mathcal{I}_+$ as part of the boundary of a bounded open set V of R^{n+1} defined by a Penrose conformal diffeomorphism from the entire spacetime.

Weak cosmic censorship conjecture *(conformal null infinity formulation). The maximal Cauchy development of generic asymptotically Euclidean initial data for the Einstein equations with physically reasonable sources possesses a complete future null infinity.*

Technical difficulties arise because of the problem of the existence and smoothness of the boundary of a Penrose diagram used to define a complete null infinity for general spacetimes. Christodoulou[15] proposes a related conjecture, which would imply the weak cosmic conjecture modulo the global existence results of Christodoulou and Klainerman, and of Klainerman and Nicolo[16].

[14] Generic is taken in its usual sense of without "exceptional properties".

[15] Christodoulou, D. (1999) *Class. Quant. Grav.*, **13**, A23–A35.

[16] Klainerman, S. and Nicolo, F. (2004) *The Evolution Problem in General Relativity*, Birkhäuser.

Weak cosmic censorship conjecture *(Christodoulou). Let (V, g) be the maximal globally hyperbolic Einsteinian development, with physically reasonable sources, of generic asymptotically Euclidean initial data, $(M, \bar{g}, K)$ and let K be a compact set on M, which is the trace of a* ***TIP*** *(terminal indecomposable past set*[17]*) γ without end point, i.e. a subset of V of the form $I^-(\gamma)$.) in V. Then the domain of dependence Ω of any open domain ω in M containing K contains a trapped surface.*

So far the only rigorous mathematical proof of the weak cosmic conjecture, due to Christodoulou, has been obtained for the spherically symmetric solutions of the Einstein equations with scalar field source.

8 Spherically symmetric Einstein scalar equations

The Einstein equations with a scalar field source with zero potential have been studied in depth by Christodoulou[18] in the case of spherical symmetry, in three space dimensions. He has obtained complete results which support the weak cosmic censorship conjecture. His results have been confirmed, and sometimes foreseen, by the numerical work of Choptuik[19].

8.1 *Spherically symmetric spacetimes*

We have given in Chapter 4 a definition of a spherically symmetric spacetime (V, g); it admits on V a metric of the form

$$g = -g_\Sigma + r^2(d\theta^2 + sin^2\theta d\varphi^2), \tag{8.1}$$

where g_Σ is a Lorentzian metric on a two-dimensional manifold Σ quotient of $V \subset R^3 \times R$ under the action of the $SO(3)$ group of rotations centred at the origin $r = 0$ of the Euclidean space R^3. The function $4\pi r^2$ is the area of the orbits, S^2 spheres, of the $SO(3)$ isometries of g. We have seen that vacuum spherically symmetric spacetimes are necessarily static (Birkhoff theorem) and are the Schwarzschild spacetimes, singular for $r = 0$; that is, with our definitions defined only when $r > 0$.

In order to study the evolution of spherically symmetric spacetimes, starting from regular initial data, Christodoulou chooses the simplest non-vacuum model; that is, the Einstein equations with source a scalar field with zero potential. He considers the case where Σ is a two-dimensional manifold with boundary represented as a subset of $R_+ \times R_+$ by coordinates $u \geq 0$ and $r \geq 0$, with the metric g_Σ of the form

$$g_\Sigma = -e^{2\nu}du^2 - 2e^{\nu+\lambda}dudr, \tag{8.2}$$

[17] A terminal indecomposable past set (TIP, Penrose terminology) is the chronological past in V of a future timelike curve.

[18] In a series of papers whose references can be found in the last one: Christodoulou, D. (1999) *Ann. Math.*, **149**, 143–217.

[19] Choptuik, M. W. (1993) *Phys. Rev. Lett.*, **70**, 9.

equivalently

$$g_\Sigma = -(e^\nu du + e^\lambda dr)^2 + e^{2\lambda} dr^2, \tag{8.3}$$

where ν and λ are functions of u and r. The curves where r alone varies are outgoing radial light rays, and the subsets $u =$ constant are, in the spacetime, future light cones with vertices on the central line $r = 0$.

8.2 *Einstein equations in adapted frame*

Recall that the Einstein equations with source a free scalar field (i.e. with zero potential) are[20]

$$S_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 8\pi T_{\alpha\beta} \equiv 8\pi \partial_\alpha \phi \partial_\beta \phi, \tag{8.4}$$

$$\nabla^\alpha \partial_\alpha \phi = 0. \tag{8.5}$$

We write them in a tetrad adapted to the geometric structure we wish to consider. This tetrad is defined by two null vectors orthogonal to the group orbit, given by

$$\ell^+ := e^{-\lambda} \frac{\partial}{\partial r}, \qquad \ell^- := e^{-\nu} \frac{\partial}{\partial u} - \frac{1}{2} e^{-\lambda} \frac{\partial}{\partial r} \tag{8.6}$$

and by two (locally defined) vectors ζ_3, ζ_4 tangent to this orbit and orthogonal between themselves.

Note that the integral curves of ℓ^+ are the outgoing light rays, those of ℓ^- the incoming light rays and that $g(\ell^\circ, \ell^-) = -1$.

The dual basis is

$$\theta^+ = \frac{1}{2} e^\nu du + e^\lambda dr, \quad \theta^- = e^\nu du, \quad \theta^3 = d\theta, \quad \theta^4 = sin\theta d\phi. \tag{8.7}$$

In this basis the spacetime metric reads

$$g = -2\theta^+ \theta^- + r^2 (d\theta^2 + sin^2 \theta d\varphi^2). \tag{8.8}$$

A straightforward computation shows that, in the chosen tetrad, the only non-vanishing components of the Einstein tensor of the metric g are all independent of θ and φ,

$$S_{--} \equiv R_{--} \equiv \frac{1}{r} \left\{ -2e^{-(\nu+\lambda)} \frac{\partial \lambda}{\partial r} + \frac{1}{2} e^{-2\lambda} \left(\frac{\partial \lambda}{\partial r} + \frac{\partial \nu}{\partial r} \right) \right\} \tag{8.9}$$

$$S_{++} \equiv R_{++} \equiv 2 \frac{e^{-2\lambda}}{r} \left(\frac{\partial \lambda}{\partial r} + \frac{\partial \nu}{\partial r} \right) \tag{8.10}$$

[20] In this section we change the unit convention for the gravitational constant, reestablishing a factor 8π in front of the stress energy tensor, in order to keep track in the equations of geometric terms and terms which come from the scalar field source.

$$S_{+-} \equiv R_{33} \equiv R_{44} \equiv \frac{e^{-2\lambda}}{r}\left(\frac{\partial\lambda}{\partial r} - \frac{\partial\nu}{\partial r}\right) + \frac{1}{r^2(1-e^{-2\lambda})} \tag{8.11}$$

$$S_{33} \equiv S_{44} \equiv R_{+-} \equiv -e^{-(\nu+\lambda)}\frac{\partial^2(\nu+\lambda)}{\partial u \partial r} + e^{-2\lambda}\left\{\frac{\partial^2\nu}{\partial r^2} + \left(\frac{\partial\nu}{\partial r} - \frac{\partial\lambda}{\partial r}\right)\left(\frac{1}{r} + \frac{\partial\nu}{\partial r}\right)\right\}. \tag{8.12}$$

If the scalar field ϕ is invariant by the rotation group, i.e. is a function of u and r alone, the only non-vanishing components of the stress energy tensor are

$$T_{++} = e^{-2\lambda}\left(\frac{\partial\phi}{\partial r}\right)^2, \quad T_{--} = \left\{e^{-\nu}\frac{\partial\phi}{\partial u} - \frac{1}{2}e^{-\lambda}\frac{\partial\phi}{\partial r}\right\}^2 \tag{8.13}$$

$$T_{33} = T_{44} = -\frac{1}{2}g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi. \tag{8.14}$$

The equations to satisfy are

$$\Sigma_{\alpha\beta} := S_{\alpha\beta} - 8\pi T_{\alpha\beta} = 0, \tag{8.15}$$

together with the wave equation for ϕ. As we have seen before, this wave equation and the Bianchi identities imply the equations

$$\nabla_\alpha \Sigma^{\alpha\beta} = 0. \tag{8.16}$$

A straightforward computation shows that if the equations

$$\Sigma_{++} = 0 \quad \text{and} \quad \Sigma_{+-} = 0 \tag{8.17}$$

are satisfied, then previous identities imply that

$$\Sigma_{33} = \Sigma_{44} = 0 \tag{8.18}$$

and the equations (8.16) give the differential equation

$$\frac{\partial\Sigma_{--}}{\partial r} + 2\left(\frac{\partial\nu}{\partial r} + \frac{1}{r}\right)\Sigma_{--} = 0 \tag{8.19}$$

whose only solution regular for $r = 0$ is $\Sigma_{--} = 0$.

We conclude that the equations $\Sigma_{++} = 0$ and $\Sigma_{+-} = 0$, together with the wave equation for ϕ, are equivalent to the full set of Einstein scalar equations if regularity at the centre is assumed.

8.3 *Reduction to one integro-differential equation*

The equations $\Sigma_{++} = 0$ and Σ_{+-} are found to be respectively equivalent to

$$\frac{\partial(\nu+\lambda)}{\partial r} = 4\pi r\left(\frac{\partial\phi}{\partial r}\right)^2 \tag{8.20}$$

and

$$\frac{\partial}{\partial r}(re^{\nu-\lambda}) - \frac{1}{r}e^{\nu+\lambda} = 0. \tag{8.21}$$

We look for solutions which are asymptotically flat at null infinity; that is, we require that for each fixed u we have

$$\lim_{r=\infty} \nu = 0, \qquad \lim_{r=\infty} \lambda = 0. \tag{8.22}$$

Equation (8.20) is then equivalent to

$$\nu + \lambda = -4\pi \int_r^\infty r \left(\frac{\partial \phi}{\partial r}\right)^2 dr, \tag{8.23}$$

and the only solution of (8.21) which exists for $r \in [0, \infty)$ is

$$e^{\nu-\lambda} = \frac{1}{r}\int_0^r e^{\nu+\lambda} dr. \tag{8.24}$$

The relations (8.23) and (8.24) express the metric in terms of the scalar field ϕ which remains our only unknown, while the only equation to solve is the wave equation for ϕ.

To simplify the writing one sets

$$h := \frac{\partial(r\phi)}{\partial r}, \qquad \alpha := e^{\nu+\lambda} \tag{8.25}$$

and denotes by an overbar the mean value between 0 and r of a function of u and r:

$$\bar{f}(u, r) := \frac{1}{r}\int_0^r f(u, \rho) d\rho. \tag{8.26}$$

With these definitions we have

$$\phi = \bar{h}, \quad \alpha = exp\left\{-4\pi \int_r^\infty \frac{1}{r}(h - \bar{h})^2 dr\right\}, \quad e^{\nu-\lambda} = \bar{\alpha}. \tag{8.27}$$

A straightforward calculation shows that the wave equation for ϕ reads then as the following non-linear integro-differential evolution equation for h

$$Dh = \frac{1}{2r}(\alpha - \bar{\alpha})(h - \bar{h}), \tag{8.28}$$

where D is a differential operator along the incoming light rays, given by

$$D := e^{\nu}\ell_- \equiv \frac{\partial}{\partial u} - \frac{1}{2}\bar{\alpha}\frac{\partial}{\partial r}. \tag{8.29}$$

8.4 *Bondi mass*

A quantity which plays an important role in radiation problems is the **Bondi mass**. In our problem it appears as follows. Consider the function of u and r given in terms of the coefficients of the metric by

$$m(u,r) = \frac{r}{2}(1 - e^{-2\lambda}) \equiv \frac{r}{2}\left(1 - \frac{\bar{\alpha}}{\alpha}\right). \tag{8.30}$$

This function is non-negative because α being a decreasing function of r at each u (see 7.23) its average $\bar{\alpha}$ on r is always at most equal to the function α. Also m vanishes for $r = 0$ and is bounded above by $\frac{r}{2}$.

Using equations (8.20), (8.21), (8.23), and (8.24), one finds that

$$\frac{\partial m}{\partial r} = 2\pi \frac{\bar{\alpha}}{\alpha}(h - \bar{h})^2. \tag{8.31}$$

Thus m is a non-increasing function of r at each u, given by

$$m(u,r) = 2\pi \int_0^r \frac{\bar{\alpha}}{\alpha}(h - \bar{h})^2(u,\rho)d\rho, \tag{8.32}$$

the function $m(u,r)$ is the "**Bondi mass**" which at retarded time u is contained within the sphere of coordinate radius r.

We have seen that if regularity at the centre is assumed, the Einstein equation $\Sigma_{--} = 0$ is a consequence of previous equations; that is, the non-linear evolution equation (8.28) for h and the Bianchi identity. In terms of m the equation $\Sigma_{--} = 0$ reads

$$Dm = -4\pi \frac{r^2}{\alpha}(D\bar{h})^2. \tag{8.33}$$

We see that $Dm \leq 0$, hence m is non-increasing along the incoming light rays.

One assumes that the initial (i.e. for $u = 0$) total Bondi mass is finite; that is,

$$\lim_{r=\infty} m(0,r) := M_0 \tag{8.34}$$

exists, finite. Then, since $m(u,r)$ is a non-decreasing function of r at each u, bounded by M_0, the limit

$$\lim_{r=\infty} m(u,r) := M(u) \tag{8.35}$$

exists for each $u \geq 0$, and is the **total Bondi mass** at u of the considered solution. Equation (8.33) implies, in our context, Bondi's theorem,

$$\frac{dM(u)}{du} \leq 0. \tag{8.36}$$

8.5 *Global existence for small data*

The initial data is a cone $C_0^+ : u = 0$; this subset and the other subsets $u =$ constant> 0, are future null cones in the constructed metric which is of the form

(8.2), with λ and ν deduced from h as explained before, and a function[21]

$$h_0(r) = h(0, r), \quad 0 \leq r < \infty. \tag{8.37}$$

In this section a solution is called global if Σ is the whole domain $u \geq 0$, $r \geq 0$. Completeness is studied later.

A classical solution of Equation (8.27) is a differentiable function $h(u, r)$ whose derivatives are continuous even on the central line $r = 0$, and which satisfies Equation (8.28).

Christodoulou proves first a semi-local (i.e. in a neighbourhood of the initial light cone) existence theorem for data $h_0 \in C^1[0, \infty)$ such that, as r tends to infinity[22], $h_0(r) = O(r^{-3})$ and $\partial h_0(r) = O(r^{-4})$. Then he proves a global existence theorem with the additional assumption that the data are small in the following scale invariant[23] norm

$$||h_0|| := \inf_{a>0} \sup_{r\geq 0} \left\{ \left[1 + \frac{r}{a}\right]^3 |h_0(r)| + \left[1 + \frac{r}{a}\right]^4 \left| \frac{dh_0(r)}{dr} \right| \right\}. \tag{8.38}$$

Theorem 8.1 *The non-linear integro-differential evolution equation admits a global solution $h \in C^1([0, \infty) \times [0, \infty))$ taking the initial value $h_0 \in C^1([0, \infty))$ at $u = 0$ if h_0 is sufficiently small in the norm (8.38).*

The corresponding spacetime is future timelike and null geodesically complete and its final Bondi mass is zero:

$$M_1 := \lim_{u\to\infty} M(u) = 0. \tag{8.39}$$

The proofs of the semi-local as well as of the global existence theorems require accurate estimates and careful use of functional analysis. They are written up in detail in Christodoulou's paper[24].

An interesting, more general consequence of inequalities obtained by Christodoulou is the following.

Theorem 8.2 *Let M_1 be the final Bondi mass of a spherically symmetric Einstein scalar spacetime, with metric written in the form (8.3). Then, the timelike lines $r = r_0$ with $r_0 > 2M_1$ are complete towards the future.*

The proof is by considering the proper time element along such a line, $e^{\nu(u,r_0)}du$. Using previous inequalities one shows that if $r_0 > 2M_1$ there exists $u_1 > 0$ such that for all $u > u_1$

$$e^{\nu(u,r_0)} \geq \bar{\alpha}(u, r_0) \geq 1 - \frac{2M(u)}{r_0} \geq \frac{1}{2}\left\{1 - \frac{2M_1}{r_0}\right\}. \tag{8.40}$$

[21] Remark that there is no constraint on h_0, except its differentiability at the vertex of the characteristic cone supporting this initial data.

[22] These conditions are the fastest fall-off preserved by evolution.

[23] The functions $h(r)$ and $h(cr)$ have the same $||.||$ norm. The choice is motivated by the scale invariance of Equation (8.28): if $h(u, r)$ is a solution, then $h(cu, cr)$ is also a solution.

[24] Christodoulou, D. (1986) *Commun. Math. Phys.*, **105**, 337–61.

Therefore, if $u_2 > u_1$

$$\int_{u_1}^{u_2} e^{\nu(u,r_0)} du \geq \frac{1}{2}\left\{1 - \frac{2M_1}{r_0}\right\}(u_2 - u_1). \tag{8.41}$$

This proper time tends to infinity when u_2 tends to infinity.

8.6 *Existence of a global generalized solution for large data*

For linear equations the theory of weak derivatives and generalized functions, also called distributions, leads easily to the definition of generalized solutions[25]. For non-linear equations the definition of a generalized solutions is a delicate one and depends on the problem at hand. Starting with the fundamental work of Leray in 1934 on the Navier–Stokes equations, various generalized solutions have been introduced in classical hydrodynamics and kinetic theory. Christodoulou introduced a definition of a generalized solution adapted to the mathematical model considered in this section.

This mathematical problem reduces to one time, one space dimension equations[26], due to the spherical symmetry. However, as in other spherically symmetric problems, the reduction eventually introduces singularities at the centre of symmetry. We have already pointed out that the Einstein scalar equation $\Sigma_{--} = 0$ is a consequence of the other equations only if regularity at the centre is assumed. If it is not assumed, one must consider this equation in addition to (8.28). We have seen that it is equivalent to the mass equation (8.33).

Christodoulou considers the Cauchy problem for the system (8.28), (8.33) with initial data on the null cone $u = 0$, $h_0 \in C^1[0, \infty)$ and such that the initial Bondi mass M_0 is finite. He denotes by Q the complement of the central line:

$$Q := \{(u, r), 0 \leq u < \infty,\ 0 < r < \infty\} \tag{8.42}$$

and defines as follows a global generalized solution of Equations (8.28) and (8.33). For $u > 0$ this solution may not be C^1 up to the central line $r = 0$.

Note that the central line $r = 0$ is a subset of zero Lebesgue measure of the domain of R^4 defined by $u \geq 0$, the domain spanned by future light cones of the metric (8.2) as long as this metric is Lorentzian.

Definition 8.3 *A global generalized solution of the considered Cauchy problem is a function h,*

$$h : (u, r) \mapsto h(u, r), \quad \textit{with} \quad h \in C^1(Q) \cap \{C^0[0, \infty), L^2(0, \infty)\}, \tag{8.43}$$

[25] Distributions are weakly infinitely differentiable, and form a ring over the set of C^∞ functions, but they cannot be multiplied, hence cannot in general be used in non-linear problems.

[26] Hyperbolic (or parabolic) partial differential equations with one space dimension have a better chance of having global classical solutions (examples are wave maps and Navier–Stokes equations) than equations in higher space dimensions, due to better embedding and multiplication properties of Sobolev spaces in lower dimensions.

such that $\bar{h}$, α, $\bar{\alpha}$ *are continuous in* $\bar{Q}$, *i.e. up to* $r = 0$, $h(0, r)$ *denoted* $h_0(r)$ *is given, and*

1. h *and* m *satisfy in* Q *the differential equations*

$$Dh = \frac{1}{2r}(\alpha - \bar{\alpha})(h - \bar{h}), \quad \textit{and} \quad Dm = -\frac{1}{2}r^2\frac{|D\bar{h}|^2}{\alpha}, \quad \textit{with} \quad m := \frac{r}{2}\left(1 - \frac{\bar{\alpha}}{\alpha}\right). \tag{8.44}$$

2. *At each* u, $\frac{\alpha}{\bar{\alpha}} \in L^1(0, r_0)$, r_0 *arbitrary and*

$$\xi := \lim_{\delta > 0} \int_\delta^r \bar{\alpha}(h - \bar{h}) r^{-1} dr \quad \textit{exists for almost all } u, \tag{8.45}$$

and is such that $r^{-1}\bar{\alpha}^{-2}\alpha\xi^2$ *is integrable over every domain* $Q_1 \subset Q$ *defined by*

$$Q_1 := \{(u, r);\ 0 < r < r_1,\ 0 < u < u_1\}, \quad r_1, u_1 \textit{ arbitrary.} \tag{8.46}$$

3. *In the sense of weak derivatives it holds that*

$$D\bar{h} = \frac{\xi}{2r}, \quad Dm = -\pi\alpha^{-1}\xi^2. \tag{8.47}$$

4. *For each pair* $(u_1, r_1) \in Q$ *the following integral identity holds*

$$\int_0^{\chi(0,r_1)} \frac{\alpha}{\bar{\alpha}}(0, r) dr = \int_0^{r_1} \frac{\alpha}{\bar{\alpha}}(u_1, r) dr + \frac{1}{2}\int_0^{u_1} \alpha(u, 0) du + 2\pi \int_0^{u_1} \int_0^{\chi(u_1,r_1)} \frac{\alpha\xi^2}{\bar{\alpha}^2 r} dr du, \tag{8.48}$$

where $\chi(u, r)$ *denotes the radial incoming light ray issued from a point* (u, r).

This integral identity, which could be called the radiation balance identity, equates the integral over the initial light cone to the sum of three positive integrals: one on the future light cone, one on the central line and a spacetime integral. All the four conditions listed above are satisfied by smooth solutions in the domain $u \geq 0$ of R^4; they are required to hold for generalized solutions.

Theorem 8.4 *For each initial data* $h_0 \in C^1[0, \infty)$ *such that at infinity* $h_0 = O(r^{-3})$, $dh_0/dr = (r^{-4})$, *there exists at least one global generalized solution of the Cauchy problem.*

Proof We give a sketch of the proof[27] It rests on constructing a solution h_ε of an approximate regular equation in the two-dimensional domain $\bar{Q}$ with boundaries, $u \geq 0$, $r \geq 0$, obtaining ε-independent estimates, and using a compactness argument to prove the existence of a (possibly non-unique) generalized solution of the original problem in the whole domain of R^4 corresponding to $u \geq 0$.

[27] Details can be found in Christodoulou, D. (1986) *Commun. Math. Phys.*, **106**, 587–621.

The approximate equation with regular coefficients, bearing on an unknown denoted h_ε which depends on u, r and a parameter ε, is

$$D_\varepsilon h_\varepsilon = \frac{1}{2(r+\varepsilon)}(\alpha_\varepsilon - \bar{\alpha}_\varepsilon)(h_\varepsilon - \bar{h}_\varepsilon), \tag{8.49}$$

where for a continuous function f_ε depending on u, r and ε one sets

$$\bar{f}_\varepsilon(u,r) := \frac{1}{r+\varepsilon}\int_0^r f_\varepsilon(u,\rho)d\rho. \tag{8.50}$$

Then elementary calculus gives

$$\bar{f}_\varepsilon(u,0) = 0 \quad \text{and} \quad \frac{\partial \bar{f}_\varepsilon}{\partial r} = \frac{f_\varepsilon - \bar{f}_\varepsilon}{\partial r}. \tag{8.51}$$

We set

$$\alpha_\varepsilon = exp\left\{-4\pi\int_r^\infty \frac{1}{r+\varepsilon}(h_\varepsilon - \bar{h}_\varepsilon)^2 dr\right\}. \tag{8.52}$$

The ε-regularized equation is

$$D_\varepsilon h_\varepsilon = \frac{1}{2(r+\varepsilon)}(\alpha_\varepsilon - \bar{\alpha}_\varepsilon)(h_\varepsilon - \bar{h}_\varepsilon), \tag{8.53}$$

where D_ε is the differential operator

$$D_\varepsilon := \frac{\partial}{\partial u} - \frac{1}{2}\bar{\alpha}_\varepsilon\frac{\partial}{\partial r}. \tag{8.54}$$

The local ε-mass function is defined by

$$m_\varepsilon := \frac{1+\varepsilon}{2}\left(1 - \frac{\bar{\alpha}_\varepsilon}{\alpha_\varepsilon}\right). \tag{8.55}$$

It is straightforward, though a bit lengthy, to show using previous definitions and the equation satisfied by h_ε, that m_ε satisfies the evolution equation

$$D_\varepsilon m_\varepsilon = -\frac{\pi}{\alpha_\varepsilon}\xi_\varepsilon^2 \quad \text{with} \quad \xi_\varepsilon := \int_0^r \bar{\alpha}_\varepsilon(h_\varepsilon - \bar{h}_\varepsilon)\frac{dr}{r+\varepsilon}. \tag{8.56}$$

The relation between $\alpha_\varepsilon/\bar{\alpha}_\varepsilon$ and m_ε leads after integration on a domain limited by an "incoming ε-light ray", i.e. by a characteristic of the first-order differential operator D_ε, to the approximate radiation balance identity,

$$\begin{aligned}\int_0^{\chi_\varepsilon(0,r_1)} \frac{\alpha_\varepsilon}{\bar{\alpha}_\varepsilon}(0,r)dr = \int_0^{r_1} \frac{\alpha_\varepsilon}{\bar{\alpha}_\varepsilon}(u_1,r)dr + \frac{1}{2}\int_0^{u_1}\alpha_\varepsilon(u,0)du \\ + 2\pi\int_0^{u_1}\int_0^{\chi_\varepsilon(u_1,r_1)} \frac{\alpha_\varepsilon\xi^2}{\bar{\alpha}^2(r+\varepsilon)}drdu,\end{aligned} \tag{8.57}$$

where $\chi_\varepsilon(u,r)$ denotes the characteristic of the operator D_ε issued from a point (u,r).

Christodoulou proves that the ε-regularized equation admits a unique global classical solution $h_\varepsilon \in C^1([0,\infty)\times[0,\infty))$ taking as given data $h_\varepsilon(0,r) \in C^1([0,\infty))$ which falls off[28] at infinity like $O(r^{-3})$ while its derivative falls off like $O(r^{-4})$. The fall-off is conserved by evolution.

One checks that the spacetime with manifold $R^+ \times R^+ \times S^2$ represented by $u \geq 0, r \geq 0, (\theta,\phi) \in S^2$, and metric given by

$$g_\varepsilon := -\alpha_\varepsilon \bar{\alpha}_\varepsilon du^2 - 2\alpha_\varepsilon du dr + (r+\varepsilon)^2(d\theta^2 + sin^2\theta), \tag{8.58}$$

together with the scalar field $\psi_\varepsilon = \bar{h}_\varepsilon$ is a classical solution of the Einstein scalar equations.

Note that the boundary $R^+ \times \{r=0\} \times S^2$ in this manifold is not a line, but a cylinder of cross-sectional area $4\pi\varepsilon^2$, which is a null hypersurface, because $\bar{\alpha}_\varepsilon(u,0) = 0$. It is an anti-event horizon – the ε-regularized problem is the original problem in the presence of a white hole[29] of a mass $m_\varepsilon = \varepsilon/2$ and boundary condition that the scalar field vanishes on the antihorizon: $\bar{h}_\varepsilon(u,0) = 0$.

Christodoulou obtains ε-independent estimates in a series of lemmas, relying to a large extent on the integral identity. He shows in particular equicontinuity of families of functions $\partial h_\varepsilon/\partial r$ and $\bar{h}_\varepsilon$. He achieves the proof of the theorem by compactness arguments and a new series of delicate estimates. □

The theorem still holds if the fall-off hypotheses on the initial data h_0 are replaced by the assumption of the finiteness of the initial Bondi mass.

A generalized solution is not necessarily unique in all its domain of existence, but it coincides with the classical solution having the same initial data in the domain of existence of the latter[30].

8.7 *Structure of generalized solutions*

The functions α and $\bar{\alpha}$ appearing in the metric of a generalized solution are continuous up to $r = 0$. This property is not enough to guarantee either uniqueness or the existence of characteristics (incoming light rays) through a point on the central line $r = 0$. However, the integral with respect to u of $\alpha(u,0)$ on a bounded interval of the central line exists, defining the proper time duration of this interval. Note that the function $\alpha(u,0)$ may vanish for a set of values of u of non-zero measure. Such a set corresponds to a set of events whose proper time duration is zero, but whose duration as observed from another line $r = constant$, in particular from infinity, is non-zero.

Christodoulou calls **regular** a **null** cone $u = u_1$ such that the integral (which appears in the radiation balance identity)

$$\int_0^{r_0} \frac{\alpha\xi^2}{\bar{\alpha}^2 r}(u_1, r)dr \tag{8.59}$$

[28] Christodoulou remarks that the result extends only when one assumes the boundedness of the initial Bondi mass.

[29] Remark due to A. Ashtekar.

[30] Christodoulou, D. (1987) *Commun. Math. Phys.*, **109**, 591–611.

exists for some r_0 (hence for all). Null cones with vertices on the central lines are regular almost everywhere in the following sense: the parameters u determining singular (i.e. non-regular) such null cones form a set of zero Lebesgue measure in R since the quantity (8.59) is integrable by point 4 of the definition of a generalized solution.

Further estimates, made with physical insight and mathematical rigour, led Christodoulou to the estimate of the total Bondi mass (Definition 8.32). He shows that

$$M(u) = M(0) - \pi \int_0^u \Xi^2(v)dv, \tag{8.60}$$

where $\pi\Xi^2(u)$ represents the radiative power at retarded time u, it may blow up at singular cones, but the integral remains defined.

With the fall off at infinity assumed from the initial data h, the total Bondi mass $M(u)$ tends to a finite, non-negative limit, as u tends to infinity, known as the **final Bondi mass**:

$$M_1 := \lim_{u=\infty} \lim_{r=\infty} \frac{1}{2} r(1 - e^{-2\lambda(u,r)}) = 2\pi \lim_{u=\infty} \int_0^\infty \frac{\bar{\alpha}}{\alpha}(h - \bar{h})^2(u, \rho)d\rho. \tag{8.61}$$

8.8 *Formation of a black hole. Cosmic censorship*

In the second of his 1987 papers Christodoulou[31] proves some more asymptotic properties of his solutions. He shows that if $M_1 > 0$, then for any constant $c > 2M_1$ the mass remaining outside the sphere of radius $r = c$, $M(u) - m(u, c)$, tends to zero as u tends to infinity. This leads to the conclusion that when u tends to infinity with fixed r:

- $\alpha(u, r)$ tends to 1 if $r > 2M_1$, and the timelike curves $r = c$ are complete if $c > 2M_1$
- $\alpha(u, r)$ tends to 0 if $r < 2M_1$, and the timelike curves $r = c$ are incomplete if $c < 2M_1$, with the length being a continuous function of c, increasing to ∞ as c tends to $2M_1$.

Christodoulou shows that an event horizon can be attached to the region corresponding to $r < 2M_1$ in $\mathcal{U}$ as a future boundary. However, this does not quite solve the cosmic censorship problem, because the generalized solution allows singular points on the central world line and is not necessarily unique.

In later articles[32] Christodoulou works with solutions of bounded variation, which he shows to exist. Singularities always emerge from the central line. Christodoulou took advantage of the fact that the conformal geometry of the two-dimensional[33] manifold Q is that of a domain in the two-dimensional Minkowski spacetime to attach a future boundary to the maximal future developments

[31] Christodoulou, D. (1987) *Commun. Math. Phys.*, **109**, 613–47.
[32] See Christodoulou, D. (1999) *Class. Quant. Grav.*, **16A**, 23–35 and references therein.
[33] Every two-dimensional pseudo-Riemannian manifold is locally conformally flat.

which are incomplete: he introduces a pair of null coordinates u, v in Q, constant respectively on the outgoing and incoming null curves and increasing towards the future. Then Q is represented as a domain D_Q in the (u, v) plane with metric given by

$$g_Q = -\Omega^2 du dv. \tag{8.62}$$

The future boundary $\mathcal{B}$ of the considered maximal future development is the future causal boundary of D_Q in Minkowski spacetime M^2. Christodoulou denotes by (u_0, v_0) the coordinates of a singular point on the central line (i.e. the end point P attached to this line); he shows that $\mathcal{B}$ contains a component $\mathcal{B}_0$, which he calls central, represented by a complete outgoing null segment

$$\mathcal{B}_0\{(u_0, v); v \in [v_0, v^*], \quad v^* \text{ a constant}; \tag{8.63}$$

this segment is replaced by a complete null curve if $\mathcal{B} \equiv \mathcal{B}_0$. If $\mathcal{B} - \mathcal{B}_0$ is not empty, it is represented by a strictly spacelike C^1 curve. Christodoulou proves several other interesting results. He gives examples where $\mathcal{B}_0$ is a complete outgoing null curve and the point P is a naked singularity. He also gives examples where $\mathcal{B}_0$ can be considered as a singular future light cone which has collapsed to a line.

8.9 *Numerical results*

The existence and stability of naked singularities in the spherically symmetric evolution of solutions of the Einstein scalar equations have been studied in depth by Choptuik[34] by numerical methods, in the context of the characteristic initial value problem considered before. He considers various 1-parameter families of initial data, such that for small values of the parameter the initial data is small, the classical solution exists globally, and the radiation energy disperses to infinity; that is, the final Bondi mass is zero. For large values of the parameter the initial data is large and the scalar field collapses into a Schwarzschild black hole. Choptuik then tuned his parameter to study the solutions, called borderline solutions, where the collapse first occurs. He discovered that these solutions possess remarkable properties; in particular his numerical investigations indicated that the borderline solutions, and only those, possess naked singularities. Since the borderline solutions are non-generic, their naked singularities do not disprove the weak cosmic censorship hypothesis.

8.10 *Instability of naked singularities*

The numerical results of Choptuik point to the instability of a naked singularity under a small change in the initial data. A rigorous analytic proof was given by Christodoulou[35]. He shows, in a sense which he explains in detail, that initial data leading to spacetimes with naked singularities are non-generic, and hence the appearance of naked singularities is an unstable phenomenon.

[34] Choptuik, M. W. (1993) *Phys. Rev. Lett.*, **70**, 9.
[35] Christodoulou, D. (1999) *Ann. Math.*, **149**, 183–217.

9 Cosmological singularities. BKL conjecture

Contributed by Thibault Damour

One usually gives the name "cosmological singularity" to the initial singularity occurring in the past of an Einsteinian description of our global Universe ("Big Bang singularity"). However, one expects that singularities of the same (local) structure will occur both in an eventual future (global) "Big Crunch singularity", and within the future developments of trapped regions ("Local Big Crunch singularities"). The singularity theorems expounded in the previous chapter do not give any information on the structure of these singularities. In contrast Belinskii, Khalatnikov, and Lifshitz[36] (BKL) proposed a rather precise description of a generic inhomogeneous cosmological singularity (in a four-dimensional vacuum Einsteinian spacetime) as being a "chaotic"[37] spacelike curvature singularity, asymptotically behaving, at each point of space, like a Bianchi IX (or VIII) homogeneous cosmological model. More precisely, BKL conjectured that, in a Gauss coordinate system ($g_{00} = -1$, $g_{0i} = 0$ for $i = 1, 2, 3$) adapted to the singularity ("located" at proper time $t = 0$), the spatial metric could be written as

$$g_{ij}(t,x) = a^2\,\ell_i\,\ell_j + b^2\,m_i\,m_j + c^2\,n_i\,n_j, \tag{9.1}$$

where the three "scale factors" (with respect to the spatial co-frame ($\ell_i\,dx^i$, $m_i\,dx^i, n_i\,dx^i$)) $a(t,x)$, $b(t,x)$, $c(t,x)$, depend mostly on time as one approaches the singularity at $t = 0$ ($\partial_t\,a \gg \partial_x\,a$, etc.) while the components $\ell_i(t,x)$, $m_i(t,x)$, $n_i(t,x)$, of the spatial co-frame depend mostly on space in this limit ($\partial_t\,\ell_i \ll \partial_x\,\ell_i$, etc.). They were able to deduce from these (approximate) assumptions a rather detailed picture of the complicated dynamics of the spatial metric in the near-singularity limit $t \to 0$. In particular, they found that, if one introduces in lieu of the three scale factors a, b, c their logarithms, α, β, γ, with

$$a = e^{\alpha}\,, \qquad b = e^{\beta}\,, \qquad c = e^{\gamma}\,, \tag{9.2}$$

and replaces the proper time t (with $g_{tt} = -1$) by a new time coordinate τ, tending to $+\infty$ as $t \to 0$, and satisfying

$$|dt| = abc\,d\tau, \tag{9.3}$$

the asymptotic behaviour of α, β, γ near the singularity is equivalent to the dynamics (with τ as time parameter) of a "particle", with coordinates $\alpha(\tau)$, $\beta(\tau)$, $\gamma(\tau)$, and free "kinetic energy" $-2(\dot\alpha\,\dot\beta + \dot\beta\,\dot\gamma + \dot\gamma\,\dot\alpha)$ moving under the influence of the superposition of several exponential potential wells, whose most important ones, in the near-singularity limit $\tau \to +\infty$, are $+\frac{1}{2}(\lambda^2\,e^{4\alpha} + \mu^2\,e^{4\beta} + \nu^2\,e^{4\gamma})$, where $\lambda = C^{\ell}_{mn}$, $\mu = C^{m}_{n\ell}$, $\nu = C^{n}_{\ell m}$ are structure constants[38] of the spatial co-frame ($\ell = \ell_i\,dx^i$, $m = m_i\,dx^i$, $n = n_i\,dx^i$). They also argued that the latter

[36] Belinskii, V. A., Lifshitz, E. M., and Khalatnikov, I. M. (1971) *Sov. Phys. Usp.*, **13**, 745; and (1972) *Sov. Phys. JETP*, **35**, 838.

[37] The BKL behaviour called here "chaotic" is often called "oscillatory" in the literature.

[38] $d\ell = C^{\ell}_{mn}\,m_\Lambda\,n$, etc.

"particle dynamics" could be well approximated, as $\tau \to +\infty$, by the simpler dynamics of a particle, with kinetic energy $-2(\dot\alpha\,\dot\beta + \dot\beta\,\dot\gamma + \dot\gamma\,\dot\alpha)$, moving within the (Lorentzian) "billiard table" defined as the intersection of the three half-spaces $\alpha \leq 0$, $\beta \leq 0$, $\gamma \leq 0$. (The "billiard walls" $\alpha = 0$, $\beta = 0$, $\gamma = 0$ are best seen as hyperplanes in the auxiliary Minkowski space $({}^3, G)$, endowed with the Lorentzian metric $G = -2(d\alpha\,d\beta + d\beta\,d\gamma + d\gamma\,d\alpha)$.)

Many studies of the BKL picture have focussed on the *a priori* much simpler case of *homogeneous* cosmological spacetimes, notably of Bianchi types VIII and IX, whose dynamics is quite similar to the one just recalled (with the drastic simplification that one is now dealing with ODEs instead of PDEs). However, even in this case there are very few mathematical results available[39] and they are far less precise than the complicated "stochastic" structures which have been delineated by physicists as taking place both in type IX Bianchi models and, at a generic point of space, in inhomogeneous BKL-type singularities[40].

On the other hand, the BKL conjecture has been extended well beyond the original framework (vacuum Einstein equations in $3 + 1$ dimensions) where it was conceived, thereby leading to surprising (heuristic) findings. When considering the vacuum Einstein equations in spacetime dimension $D \equiv d + 1$, it was found that they generically[41] behave, near a spacelike singularity, in a "chaotic" BKL-type manner for spacetime dimensions $D \leq 10$, while in higher spacetime dimensions $D \geq 11$, the asymptotic behaviour near the singularity drastically simplifies, to become monotonic and power-law-like[42]. Interestingly this shows that the Einstein vacuum equations implicitly "know" that the spacetime dimensions $D = 10$ and $D = 11$ are special[43].

When considering more general Einstein-matter systems, both types of behaviour ("chaotic" versus "monotonic power-law") can occur, depending on the type of matter fields that one considers. For instance, the Einstein scalar field system was heuristically predicted[44] to exhibit a "non-chaotic" power-law behaviour. As we shall discuss in detail (in some examples) in the next section, a precise mathematical formulation of such non-chaotic asymptotic power-law behaviours near spacelike singularities has been developed ("Asymptotically Velocity Term Dominated", AVTD, behaviour), and results on Fuchsian systems have been used to prove theorems concerning, for instance, the Einstein scalar

[39] See Ringström, H. (2001) *Ann. H. Poincaré*, **2**, 405–500.

[40] See notably Khalatnikov, I. M., Lifshitz, E. M., Khanin, K. M., Shchur, L. N., and Sinai, Ya. G. (1985) *J. Stat. Phys.*, **38**, 97–114.

[41] In the context of the BKL conjecture, one loosely qualifies as "generic" a purported (or approximate) solution of the field equations whose initial data contain the maximum allowed number of free spatial functions, given the constraints restricting Cauchy data.

[42] Demaret, J., Henneaux, M., and Spindel, P. (1985) *Phys. Lett. B*, **164**, 27; Demaret, J., Hanquin, J. L., Henneaux, M., Spindel, P., and Taormina, A. (1986) *Phys. Lett. B*, **175**, 129.

[43] We recall that superstring theories can be consistently formulated only in spacetime dimension $D = 10$, while maximal supergravity theory lives in spacetime dimension $D = 11$.

[44] Belinskii, V. A. and Khalatnikov, I. M. (1973) *Sov. Phys. JETP*, **36**, 591–7.

system[45], or the vacuum Einstein equations in spacetime dimensions $D \geq 11$[46]. On the other hand, many Einstein-matter systems of interest for physics have been (heuristically) found to exhibit a chaotic behaviour similar to the standard BKL picture recalled above. This is the case, notably, for the Einstein–Maxwell system in any dimension $D \geq 4$[47], and for the specific Einstein scalar p-form systems describing the low-energy limits of the (bosonic sectors of the) various $D = 10$ superstring theories, as well as the Einstein three-form system in $D = 11$ describing the bosonic sector of maximal supergravity[48].

The alert reader will have noticed that our brief presentation above of the original BKL conjecture was somewhat vague (even by physicists' standards) and certainly not mathematically precise. The main problems in mathematically formulating the conjecture are: (i) to define, in precise terms, the meaning of the decomposition (9.1) of the spatial metric into three scale factors a, b, c, and three covectors ℓ_i, m_i, n_i, and (ii) to characterize in what precise sense the scale factors are "mainly time-dependent", while the covectors ℓ_i, m_i, n_i are "mainly space-dependent". Contrary to the simpler case of non-chaotic systems (such as the Einstein-scalar case) it is not possible to describe the chaotic BKL behaviour by neglecting, in Einstein equations, all the terms involving spatial derivatives[49]. Indeed, the chaotic BKL behaviour for vacuum gravity (though not for the systems mentioned in the previous footnote) crucially depends on considering both the time-dependence of the scale factors, and the space-dependence of the coframe. Some progress on a more precise formulation of the chaotic BKL behaviour has been obtained in simplified situations, involving two, or even only one, isometries. A combination of analytic approximations and of numerical investigations of certain symmetry-specialized classes of inhomogeneous cosmological spacetimes has given good numerical evidence for the validity of the (generalized) BKL conjecture[50]. Recently, some numerical work[51] has also confirmed some aspects of the BKL conjecture in absence of any isometries.

To end this section, let us briefly summarize another line of work[52] which seems to provide both new (conjectural) results concerning the chaotic BKL behaviour of very general systems, and a mathematically precise formulation of

[45] Andersson, L. and Rendall, A. (2001) *Commun. Math. Phys.*, **218**, 479–511.

[46] Damour, T., Henneaux, M., Rendall, A., and Weaver, M. (2002) *Ann. H. Poincaré*, **3**, 1049.

[47] Damour, T. and Henneaux, M. (2000) *Phys. Lett. B*, **488**, 108.

[48] Damour, T. and Henneaux, M. (2000) *Phys. Rev. Lett.*, **85**, 920.

[49] Note, however, that the generic chaotic behaviour of certain inhomogeneous systems (such as the Einstein–Maxwell system in $D \geq 4$, or the Einstein 3-form system in $D \geq 7$) *can* be captured by retaining only the terms involving only time derivatives. This shows that the implication, non-chaotic $\Rightarrow$ AVTD, cannot be reversed.

[50] Weaver, M., Isenberg, J., and Berger, B. K. (1998) *Phys. Rev. Lett.*, **80**, 2984; Berger, B. K. and Moncrief, V. (1998) *Phys. Rev. D*, **58**, 064023; Berger, B. K., Isenberg, J., and Weaver, M. (2004) *Phys. Rev. D*, **64**, 84–6.

[51] Garfinkle, D. (2007) *Class. Quant. Grav.*, **12**, S295–S306 and references therein.

[52] Damour, T., Henneaux, M., and Nicolai, H. (2003) *Class. Quant. Grav.*, **20**, R145–R200; Damour, T. and de Buyl, S. to be published.

the (generalized) BKL conjecture. This line of work uses several technical tools, notably:

- a quasi-Gaussian coordinate system defined by a vanishing "shift" (i.e. $g_{0i} = 0$), and a metric of the form ($i, j = 1, \ldots, d$ when working in $D = d + 1$ dimensions)

$$ds^2 = -(\tilde{N}\sqrt{g}\, d\tau)^2 + g_{ij}\omega^i\, \omega^j\,, \tag{9.4}$$

  where $g \equiv \det g_{ij}$ and where the "rescaled lapse" $\tilde{N}$ is required to be equal to unity: $\tilde{N} = 1$[53];
- the replacement of the $d(d+1)/2$ spatial metric components $g_{ij}(\tau, x^m)$ by a new set of metric variables: d "diagonal degrees of freedom" $\beta^a(\tau, x^m)$, $a = 1, \ldots, d$; together with $d(d-1)/2$ "off-diagonal degrees of freedom" $\mathcal{N}^a_i(\tau, x^m)$. These new variables are (spacetime) local (analytic) functions of the g_{ij}'s which are uniquely defined by requiring that the (positive-definite) spatial metric components read

$$g_{ij} = \sum_{a=1}^{d} e^{-2\beta^a}\, \mathcal{N}^a_i\, \mathcal{N}^a_j\,, \tag{9.5}$$

  where $\mathcal{N}^a_i$ is required to be an *upper triangular* matrix ($\mathcal{N}^a_i = 0$, if $i < a$) with ones on the diagonal ($\mathcal{N}^a_i = 1$ when $i = a$); Equation 9.5 is referred to as an "**Iwasawa decomposition**" of g_{ij}, and thereby β^a, $\mathcal{N}^a_i$ as "Iwasawa variables";
- the use of a Hamiltonian formalism (à la Arnowitt–Deser–Misner) which associates to the Iwasawa variables $\beta^a(\tau, x)$, $^a_i(\tau, x)$, their respective "conjugate momenta", $\pi_a(\tau, x)$, $\mathcal{P}^i_a(\tau, x)$ (the latter being defined only for $i > a$), and allows one to work with a first-order-in-time evolution system for the phase-space variables ($\beta(\tau)$, $\mathcal{N}(\tau)$, $\pi(\tau)$, $\mathcal{P}(\tau)$)[54].

  With this notation in hand, one can state the following conjectural findings which can be thought of as precise versions of the assumptions (such as $\partial_t\, a \gg \partial_x\, a$, $\partial_t\, \ell_i \ll \partial_x\, \ell_i$, etc.) made in the original BKL approach.
- When the coordinate time τ of (9.4) tends to $+\infty$ (which corresponds to the near-singularity limit, "$t \to 0$") the "off-diagonal Iwasawa degrees of freedom" $\mathcal{N}^a_i(\tau, x)$ and $\mathcal{N}^i_a(\tau, x)$, considered at any given spatial point x, have *limits*[55], say (suppressing indices) $\mathcal{N}_{(0)}(x) = \lim_{\tau\to+\infty} \mathcal{N}(\tau, x)$, and $\mathcal{P}_{(0)}(x) = \lim_{\tau\to+\infty} \mathcal{P}(\tau, x)$.

[53] Here, we consider a spacetime $M_d \times$ and a time-independent co-frame $\omega^i(x) = \omega^i_m(x^n)\, dx^m$ on the spatial d-dimensional manifold M_d, with x^m, $m = 1, \ldots, d$, denoting some (local) coordinate system on M_d.

[54] For simplicity, we consider here only the vacuum Einstein case; the general case involving scalar fields and an arbitrary menu of p-form fields is dealt with by a simple extension of the same tools, involving further phase-space variables to describe the matter degrees of freedom.

[55] These limits are expected to be finite, except when x belongs to a set of measure zero. See discussion in the papers quoted in footnote 51.

- By contrast, the "diagonal Iwasawa degrees of freedom" $\beta^a(\tau,x)$, $\pi_a(\tau,x)$ have no limits as $\tau \to +\infty$, but their asymptotic behaviour as $\tau \to +\infty$ can be described, at each given spatial point x, by a certain first-order-in-τ system of ODEs, whose coefficients are algebraic in $\mathcal{N}_{(0)}(x)$, $\mathcal{P}_{(0)}(x)$ and $\partial_x \mathcal{N}_{(0)}(x)$. More precisely, the asymptotic behaviour, at a given x, of $\beta(\tau,x)$, $\pi(\tau,x)$, is that of a solution $\beta_{(0)}(\tau,x)$, $\pi_{(0)}(\tau,x)$ of the Hamilton evolution equations following from the "asymptotic Hamiltonian" (using canonical Poisson brackets $\{\beta^a, \hat{E}\pi_b\} = \delta^a_b$)

$$\mathcal{H}^{\text{asymp}}(\beta,\pi) = \frac{1}{4}\, G^{ab}\, \pi_a\, \pi_b + \mathcal{V}_S^{\text{asymp}} + \mathcal{V}_G^{\text{asymp}}\,, \tag{9.6}$$

where

$$G^{ab}\, \pi_a\, \pi_b \equiv \sum_{a=1}^{d} \pi_a^2 - \frac{1}{d-1}\left(\sum_{a=1}^{d} \pi_a\right)^2$$

contains the "contravariant form" of the Lorentzian metric

$$G_{ab}\, \beta^a\, \beta^b = \sum_{a=1}^{d} (\beta^a)^2 - \left(\sum_{a=1}^{d} \beta^a\right)^2,$$

and where

$$\mathcal{V}_S^{\text{asymp}} = \frac{1}{2} \sum_{a=1}^{d-1} e^{-2(\beta^{a+1}-\beta^a)} (\mathcal{P}^i_{(0)a}\, \mathcal{N}^{a+1}_{(0)i})^2\,, \tag{9.7}$$

(where $i = 1, \ldots, d$ is summed over) and

$$\mathcal{V}_G^{\text{asymp}} = \frac{1}{2}\, e^{-2\alpha_{1d-1d}(\beta)} (C^1_{(0)d-1d})^2\,. \tag{9.8}$$

In the last equation, $\alpha_{abc}(\beta)$ (for $b \neq c$) denotes the linear form $\alpha_{abc}(\beta) = \beta^a + \sum_{e\neq b,c} \beta^e$ (evaluated for $a = 1$, $b = d-1$ and $c = d$), and $C^a_{(0)bc}$ (with $b \neq c$ and $C^a_{(0)bc} = -C^a_{(0)cb}$) denote the structure functions ($d\theta^a_{(0)} = -\frac{1}{2}\, C^a_{(0)bc}\, \theta^b_{(0)\Lambda}\, \theta^c_{(0)}$) of the "asymptotic Iwasawa frame" $\theta^a_{(0)}(x) = \mathcal{N}^a_{(0)i}(x)\, \omega^i$. Note that all the coefficients entering the exponential potential terms[56] (9.7), (9.8) depend only on the

[56] Note that there are d exponential potential terms entering the asymptotic Hamiltonian, say $c_k \exp -2w_k(\beta)$ with $k = 1, \ldots, d$, where the d exponents $w_k(\beta)$ are *linear forms* in the βs. A remarkable fact (see references cited in the review of Damour, Henneaux, and Nicolai mentioned above) is that these d linear forms can be identified with the d simple roots of a Lorentzian Kac–Moody algebra, which is AE_d in the pure gravity case considered here. In other words, the "asymptotic billiard chamber" defined by the inequalities $w_k(\beta) \geq 0$ (within which the motion of the βs is approximately asymptotically confined) can be identified with the "Weyl chamber" of AE_d. Other Kac–Moody algebras were similarly found to enter the asymptotic dynamics of some Einstein-matter systems: for instance the ten exponential potential terms entering the asymptotic Hamiltonian for eleven-dimensional supergravity were found to be related to the "last" hyperbolic "exceptional" Kac–Moody algebra E_{10}.

spatial point (through $\mathcal{N}_{(0)}(x)$, $\mathcal{P}_{(0)}(x)$ and $\partial_x \mathcal{N}_{(0)}(x)$ which enters $C_{(0)}$), so that the asymptotic evolution system for β and π constitutes, at each point of space, a well-defined system of ODEs. When completing this asymptotic evolution system by the ODEs expressing the time-independence of $_{(0)}$ and $_{(0)}$ one gets a "chaotic analogue" of the AVTD evolution system considered in the non-chaotic, monotonic power-law case of the form:

$$\begin{aligned}
\partial_\tau \beta^a_{(0)} &= \frac{1}{2} G^{ab} \pi^{(0)}_b \,, \\
\partial_\tau \pi^{(0)}_a &= -\frac{\partial}{\partial \beta^a_{(0)}} \{ \mathcal{V}^{\text{asymp}}_S(\beta_{(0)}; \mathcal{P}_{(0)}, \mathcal{N}_{(0)}) \\
&\qquad + \mathcal{V}^{\text{asymp}}_G(\beta_{(0)}; \mathcal{N}_{(0)}, \mathcal{P}_{(0)}, \partial_x \mathcal{N}_{(0)}) \} \,, \\
\partial_\tau \mathcal{N}^a_{(0)i} &= 0 \,, \\
\partial_\tau \mathcal{P}^i_{(0)a} &= 0 \,.
\end{aligned} \tag{9.9}$$

The evolution system (9.9) must be completed by adding some "asymptotic constraints", say

$$\begin{aligned}
\mathcal{H}^{\text{asymp}}(\beta_{(0)}, \pi_{(0)}, \mathcal{N}_{(0)}, \partial_x \mathcal{N}_{(0)}, \mathcal{P}_{(0)}) &= 0 \,, \\
\mathcal{H}^{\text{asymp}}_a(_{(0)}, \partial_x {}_{(0)}, {}_{(0)}) &= 0 \,,
\end{aligned} \tag{9.10}$$

where $\mathcal{H}^{\text{asymp}}$ is the (conserved) quantity defined in (9.6), and where the definition of $\mathcal{H}^{\text{asymp}}_a$ will be found in the references cited above. Note that both constraints are preserved by the asymptotic evolution system (9.9).

Finally, this leads to the following precise formulation of the **BKL conjecture in Iwasawa variables**. Let, for $x \in U$, $(\beta_{(0)}(\tau,x)$, $\pi_{(0)}(\tau,x)$, $\mathcal{N}_{(0)}(\tau,x), \mathcal{P}_{(0)}(\tau,x))$ be a solution of the asymptotic evolution system (9.9), satisfying the asymptotic constraints (9.10), and such that the d x-dependent coefficients $\mathcal{P}_{(0)}\mathcal{N}_{(0)}$ and $C_{(0)}$ (whose squares define the coefficients of the d exponential potential terms (9.7) and (9.8)) do not vanish in the considered spatial domain. Then there exists a solution $(\beta(\tau,x)$, $\pi(\tau,x)$, $\mathcal{N}(\tau,x), \mathcal{P}(\tau,x))$ of the vacuum Einstein equations (including the constraints) such that the differences $\bar{\beta}(\tau,x) \equiv \beta(\tau,x) - \beta_{(0)}(\tau,x)$, $\bar{\pi}(\tau,x) \equiv \pi(\tau,x) - \pi_{(0)}(\tau,x)$, $\bar{\mathcal{N}}(\tau,x) \equiv \mathcal{N}(\tau,x) - \mathcal{N}_{(0)}(\tau,x)$, $\bar{\mathcal{P}}(\tau,x) \equiv \mathcal{P}(\tau,x) - \mathcal{P}_{(0)}(\tau,x)$ tend to zero as $x \in U$ is fixed and $\tau \to +\infty$.

In addition the references mentioned above give a qualitative description of the way the various differences $\bar{\beta}, \bar{\pi}, \bar{\mathcal{N}}, \bar{\mathcal{P}}$ tend to zero as $\tau \to +\infty$, and suggest a certain generalization of the Fuchsian system to describe this asymptotic behaviour.

However, for the time being no mathematical results of this type (applicable to "chaotic systems") have been proven. Therefore, in the next section we

shall consider only "non-chaotic systems"[57] whose asymptotic behaviour can be mathematically tackled by certain Fuchsian systems.

10 AVTD behaviour

10.1 *Definitions*

A particular case of asymptotic behaviour at a singularity is the case where space–space interaction becomes negligible as one approaches a singularity; namely, an Einsteinian spacetime will approach a VTD (Velocity Term Dominated) spacetime, defined as follows.

Definition 10.1 *A VTD spacetime is a spacetime which is solution of the equations obtained by dropping space derivatives from the Einstein equations.*

A VTD spacetime is solution of the full Einstein equations only if its metric coefficients depend only on time. Particular Einsteinian VTD spacetimes are the Kasner spacetimes described in Chapter 5.

Definition 10.2 *An Einsteinian spacetime is said to have AVTD (Asymptotically Velocity Term Dominated) behaviour if it approaches a VTD spacetime as it approaches a singularity.*

The four-dimensional Einsteinian vacuum spacetimes do not in general have AVTD behaviour.

The four-dimensional Einsteinian spacetimes with source a scalar field[58], and some vacuum Einsteinian spacetimes with an isometry group[59] have been proved to have AVTD behaviour.

A study for higher-dimensional spacetimes[60] has shown that the vacuum Einstein equations have AVTD behaviour in dimension $d \geq 11$.

10.2 *Fuchs theorem*

One of the tool which has proved to be very successful in showing that AVTD behaviour is found in solutions of Einstein's equations is the Fuchsian analysis.

[57] Note that a precise definition of "chaotic" versus "non-chaotic", within the present context, is best formulated by using the definition of the (fundamental) "asymptotic billiard chamber" introduced above, i.e. the region in the Lorentzian β-space defined by the inequalities $w_k(\beta) \geq 0$ for $k = 1, \ldots, d$. An asymptotic system is called "chaotic" when its asymptotic billiard chamber is contained within the future light-cone $G_{ab}\,\beta^a\,\beta^b \leq 0$, $\sum_a \beta^a \geq 0$ (where G_{ab} is the Lorentzian metric in β-space defined above). When this is not the case (because the chamber extends partially outside the light cone), the system is called "non-chaotic", and its asymptotic behaviour can be described by an AVTD system.

[58] Andersson, L. and Rendall, A. (2001) *Commun. Math. Phys.*, **218**, 479–511.

[59] Rendall, A. (2000) *Class. Quant. Grav.*, **17**, 3305–16 (Gowdy spacetime).
Isenberg, J. and Moncrief, V. (2002) *Class. Quant. Grav.*, **19**, 5361–86 (U(1) isometry, space T^3 and a half polarization condition).

[60] Damour, T., Henneaux, M., Rendall, A., and Weaver, M. (2002) *Ann. H. Poincaré*, gr-qc 0202069v1.

The idea is to write field variables as the sum of a VTD solution and a remainder field, and then one shows that the remainder field satisfies a Fuchsian type system which guarantees that it vanishes as one approaches the singularity. A discussion of Fuchsian systems appears in Appendix V.

The Fuchs theorem concerns the first-order differential system (for mathematical convenience we choose the time parameter t such that t tends to $+\infty$ when one approaches the singularity)

$$\partial_t u - A(x)u = e^{-t\mu} f(t, x, u, D_x u). \tag{10.1}$$

The unknown u is a set of tensor fields on a manifold $V = M \times R$, $x \in M$ a real analytic manifold, $t \in R$. The coefficient A is a linear operator on u continuous in t and analytic in $x \in M$, while f is a mapping analytic in $x \in M$, continuous in $t \in [0, T)$, analytic in u if $|u| \leq c$, c some number, and linear[61] in the space derivative $D_x u$; that is,

$$f(t, x, u, D_x u) \equiv f_0(t, x, u) + f_1(t, x, u) D_x u. \tag{10.2}$$

We use the following definition from Appendix V, supposing that the space manifold M admits an holomorphic[62] extension $\hat{M}$

Definition 10.3 *Equation (10.1) is called Fuchsian if*

1. *The linear operator $A(x)$ admits a holomorphic extension $A(z)$ to $\hat{M}$ and the mappings $(x, X) \mapsto f_i(t, x, X)$, $i = 0, 1$, admit holomorphic extensions, continuous in $t \in [0, T]$, to $\hat{M} \times U$, where $U := \{|X| < c\}$.*

2. *The number μ is positive, $\mu > 0$.*

3. *There exists a number $\alpha < 1$ and a number $\Sigma > 0$ such that the linear operator $\sigma^{\mu^{-1}A(z)} := e^{\mu^{-1}A(z) log \sigma}$ satisfies the inequality*

$$Sup_{z \in \hat{M}} |\sigma^{\mu^{-1}A(z)}| \sigma^{\alpha} \leq \Sigma \quad for \quad 0 \leq \sigma \leq 1. \tag{10.3}$$

We prove in the appendix that a sufficient condition for point 3 to hold is that the real parts of the eigenvalues of A are greater than $-\mu$, and we prove the following theorem.

Theorem 10.4 *If Equation (10.1) is Fuchsian, there exists a number $T > 0$ such that this equation admits a solution u analytic in $x \in M$, C^1 in $t > T$ and u tending to zero as t tends to infinity.*

[61] This property can be attained for more general systems by derivating the equations and introducing derivatives as new unknowns.

[62] That is, complex analytic.

11 Case of 1-parameter spatial isometry

We give as examples of asymptotic VTD or non-AVTD asymptotic behaviour the case of four-dimensional vacuum Einsteinian spacetimes with one-parameter spacelike isometry group on topologically general manifolds[63].

11.1 *Equations*

Let $(V_4 \equiv M \times R, \hat{g})$ be a spacetime with a spacelike isometry group which endows[64] M with the structure of a principal fibre bundle over a surface Σ. The metric $\hat{g}$ can be written

$$\hat{g} \equiv e^{-2\phi\,(3)}g + e^{2\phi}(d\theta + a)^2, \qquad {}^{(3)}g \equiv {}^{(3)}\, g_{\alpha\beta}dx^\alpha dx^\beta \tag{11.1}$$

with θ a parameter on the orbit of the isometry group, ϕ a scalar, a a locally defined 1-form and ${}^{(3)}g$ a Lorentzian metric, all on $V_3 := \Sigma \times R$.

If the orbits of the isometry group are orthogonal to 3-manifolds the 1-form a is identically zero; the case, $a \equiv 0$ is called the **polarized case**.

The vacuum 3+1 Einstein equations $Ricci(\hat{g}) = 0$ for such a metric on V_4 are known[65] to be equivalent[66] to the wave map equation from $(V_3, {}^{(3)}\, g)$ into the Poincaré plane $P =: (R^2, G)$, coupled to the 2+1 Einstein equations for ${}^{(3)}g$ on V_3 with source the wave map. The wave map $\Phi : V_3 \to R^2$ is defined by a pair of scalar functions ϕ and ω, and the metric of the Poincaré plane is

$$G \equiv 2(d\phi)^2 + \frac{1}{2}e^{-4\gamma}(d\omega)^2, \tag{11.2}$$

The scalar function ω on V_3 is linked to the differential F of a by the relation

$$d\omega = e^{4\phi} * F, \quad \text{with} \quad F = da. \tag{11.3}$$

In local coordinates x^α, $\alpha = 0, 1, 2$, on V_3, with η the volume form of ${}^{(3)}g$, this reads as

$$F_{\alpha\beta} \equiv \frac{1}{2}e^{-4\phi}\eta_{\alpha\beta\lambda}\partial^\lambda\omega. \tag{11.4}$$

The wave map equations are, with ${}^{(3)}\nabla$ the covariant derivative in the metric ${}^{(3)}g$,

$$g^{\alpha\beta}\left({}^{(3)}\nabla_\alpha\partial_\beta\phi + \frac{1}{2}e^{-4\phi}\partial_\alpha\omega\partial_\beta\omega\right) = 0 \tag{11.5}$$

$$g^{\alpha\beta}\left({}^{(3)}\nabla_\alpha\partial_\beta\omega - 4\partial_\alpha\omega\partial_\beta\phi\right) = 0. \tag{11.6}$$

[63] Choquet-Bruhat, Y., Isenberg, J., and Moncrief, V. (2005) *Il Nuovo Cimento B*, **119**(7–9) (polarized case); Choquet-Bruhat, Y. and Isenberg, J. (2006) *J. Geom. Phys.*, **6**, 1199–214 (half polarized case).

[64] See Appendix VII or Chapter 14.

[65] See Choquet-Bruhat, Y. and Moncrief, V., reference in Chapter 16.

[66] If we choose an arbitrary harmonic 1-form appearing in the solution to be zero.

In the polarized case the wave map equations reduce to the scalar wave equation

$$g^{\alpha\beta(3)}\nabla_\alpha\partial_\beta\phi = 0. \tag{11.7}$$

The $2+1$ Einstein equations are, with a dot denoting the scalar product in the metric G:

$${}^{(3)}R_{\alpha\beta} = \partial_\alpha\Phi.\partial_\beta\Phi := 2\partial_\alpha\phi\partial_\beta\phi + \frac{1}{2}e^{-4\gamma}\partial_\alpha\omega\partial_\beta\omega. \tag{11.8}$$

To solve these equations we choose for ${}^{(3)}g$ a zero shift, we denote the lapse by e^λ and we weigh by e^λ, without restricting the generality, the t-dependent space metric $g = g_{ab}dx^a dx^b$, $a, b = 1, 2$; that is, we set

$${}^{(3)}g \equiv -N^2dt^2 + g_{ab}dx^a dx^b \quad \text{with} \quad N \equiv e^\lambda, \quad g_{ab} \equiv e^\lambda\sigma_{ab}. \tag{11.9}$$

We denote by σ^{ab} the contravariant form of σ. The extrinsic curvature of Σ_t in $(V_3, {}^{(3)}g)$ is

$$k_{ab} := -\frac{1}{2N}\partial_t g_{ab} \equiv -\frac{1}{2}(\sigma_{ab}\partial_t\lambda + \partial_t\sigma_{ab}). \tag{11.10}$$

The mean extrinsic curvature τ is therefore

$$\tau := g^{ab}k_{ab} \equiv -e^{-\lambda}\left(\partial_t\lambda + \frac{1}{2}\psi\right) \tag{11.11}$$

where we have set

$$\psi =: \sigma^{ab}\partial_t\sigma_{ab}. \tag{11.12}$$

The connection coefficients (Christoffel symbols) of ${}^{(3)}g$ are found to be (note that ${}^{(3)}g^{00} = -e^{-2\lambda}$, ${}^{(3)}g^{ab} = g^{ab} = e^{-\lambda}\sigma^{ab}$) :

$${}^{(3)}\Gamma^c_{ab} = \Gamma^c_{ab}(g) = \Gamma^c_{ab}(\sigma) + \frac{1}{2}(\delta^c_b\partial_a\lambda + \delta^c_a\partial_b\lambda - \sigma^{cd}\sigma_{ab}\partial_d\lambda) \tag{11.13}$$

$${}^{(3)}\Gamma^0_{00} = \partial_t\lambda, \quad {}^{(3)}\Gamma^0_{0a} = \partial_a\lambda, \quad {}^{(3)}\Gamma^a_{00} = \sigma^{ab}e^\lambda\partial_a\lambda, \tag{11.14}$$

$${}^{(3)}\Gamma^0_{ab} = -e^{-\lambda}k_{ab}, \quad {}^{(3)}\Gamma^b_{a0} = -e^\lambda k^b_a. \tag{11.15}$$

In particular it holds that

$${}^{(3)}g^{\alpha\beta(3)}\Gamma^0_{\alpha\beta} = \frac{1}{2}\psi e^{-2\lambda}. \tag{11.16}$$

We see that the metric ${}^{(3)}g$ is in harmonic time if and only if $\psi = 0$.

The Einstein equations split into constraints and evolution equations. We denote by $S^\alpha_\beta \equiv {}^{(3)}R^\alpha_\beta - \frac{1}{2}\delta^\alpha_\beta {}^{(3)}R$ the Einstein tensor of ${}^{(3)}g$, by T^α_β the stress energy tensor of Φ, and we set $\Sigma^\alpha_\beta \equiv S^\alpha_\beta - T^\alpha_\beta$. The constraints are:

$$C_0 \equiv \Sigma^0_0 \equiv -\frac{1}{2}\{R(g) - k.k + \tau^2 - e^{-2\lambda}\partial_t\Phi.\partial_t\Phi - g^{ab}\partial_a\Phi.\partial_b\Phi\} = 0 \tag{11.17}$$

and (indices raised with g^{ab}, ∇ the covariant derivative in the metric g)

$$C_a \equiv e^{\lambda}\Sigma^0_a \equiv -\{\nabla_b k^b_a - \partial_a\tau + e^{-\lambda}\partial_a\Phi.\partial_b\Phi\} = 0. \tag{11.18}$$

The evolution equations are, with $N = e^{\lambda}$,

$$N({}^{(3)}R^b_a - \rho^b_a) \equiv -\partial_t k^b_a + N\tau k^b_a - \nabla^b\partial_a N + NR^b_a - N\partial_a\Phi.\partial^b\Phi = 0. \tag{11.19}$$

In order to obtain a first-order system in the Fuchsian analysis that we will make, we introduce new unknowns $\Phi_t, \Phi_a, \sigma^{ab}_c$ which will be identified with the first partial derivatives of Φ and the covariant derivative of σ with respect to a given, t-independent, metric $\tilde{\sigma}$. These new unknowns satisfy the equations

$$\partial_t\Phi = \Phi_t, \qquad \partial_t\Phi_a = \partial_a\Phi_t, \tag{11.20}$$

$$\partial_t\sigma^{ab}_c = \tilde{\nabla}_c\partial_t\sigma^{ab} \tag{11.21}$$

where, by the definitions of σ and k:

$$\partial_t\sigma^{ab} = 2e^{2\lambda}k^{ab} + \sigma^{ab}\partial_t\lambda. \tag{11.22}$$

The function λ will not be an unknown, but will be determined by a gauge condition from its VTD value.

11.2 *VTD solutions of the* $2+1$ *Einstein evolution equations*

The Velocity Terms Dominated equations are obtained by dropping the space derivatives in the original equations.

We denote with a tilde quantities independent of t, and we denote VTD solutions with a hat.

In order to obtain a global (on Σ) formulation we choose a VTD metric which remains in a fixed conformal class over Σ as t evolves; we set

$$\hat{\sigma}_{ab} = \tilde{\sigma}_{ab} \quad \text{and} \quad \hat{g}_{ab} = e^{\hat{\lambda}}\tilde{\sigma}_{ab}. \tag{11.23}$$

Then

$$\hat{\psi} = 0, \quad \hat{g}^{ab} = e^{-\hat{\lambda}}\tilde{\sigma}^{ab} \quad \partial_t\hat{g}_{ab} = e^{\hat{\lambda}}\tilde{\sigma}_{ab}\partial_t\hat{\lambda}, \tag{11.24}$$

and the definition of k gives that

$$\hat{k}_{ab} = -\frac{1}{2}\tilde{\sigma}_{ab}\partial_t\hat{\lambda}, \quad \hat{k}^b_a = -\frac{1}{2}e^{-\hat{\lambda}}\delta^b_a\partial_t\hat{\lambda}, \quad \hat{\tau} =: \hat{k}^a_a = -e^{-\hat{\lambda}}\partial_t\hat{\lambda}. \tag{11.25}$$

Writing that these VTD quantities satisfy the VTD evolution equations gives

$$\partial_t\hat{k}^b_a = \hat{N}\hat{\tau}\hat{k}^b_a. \tag{11.26}$$

Therefore, by straightforward computation, we have

$$\partial^2_{tt}\hat{\lambda} = 0, \qquad \text{hence} \qquad \hat{\lambda} = \tilde{\lambda} - \tilde{v}t \tag{11.27}$$

with $\tilde{\lambda}$ and $\tilde{v}$ arbitrary functions on Σ, independent of t. Then

$$\hat{k}_{ab} = \frac{1}{2}\tilde{v}\tilde{\sigma}_{ab}, \quad \hat{k}^b_a = \frac{1}{2}\tilde{v}e^{-\hat{\lambda}}\delta^b_a, \quad \hat{\tau} = e^{-\hat{\lambda}}\tilde{v}. \tag{11.28}$$

The VTD constraints reduce to

$$\hat{C}_0 \equiv \hat{\Sigma}^0_0 \equiv -\frac{1}{2}\{-\hat{k}.\hat{k} + \hat{\tau}^2 - e^{-2\hat{\lambda}}\partial_t\hat{\Phi}.\partial_t\hat{\Phi}\} = 0 \tag{11.29}$$

11.3 *The polarized case*

11.3.1 *Scalar wave and constraint VTD solution*

The VTD equation for ϕ reduces to

$$\partial^2_{tt}\hat{\phi} = 0 \quad \text{hence} \quad \partial_t\hat{\phi} = -\frac{1}{2}\tilde{w} \tag{11.30}$$

(the factor 2 is inserted to simplify the formula to come). The VTD constraint is therefore

$$-\hat{k}.\hat{k} + \hat{\tau}^2 - 2e^{-2\hat{\lambda}}\partial_t\hat{\phi}\partial_t\hat{\phi} = \frac{e^{-2\lambda}}{2}(\tilde{v}^2 - w^2) = 0. \tag{11.31}$$

We choose

$$\tilde{v} = \tilde{w}; \quad \text{hence} \quad \hat{\phi} = \tilde{\phi} - \frac{1}{2}\tilde{v}t. \tag{11.32}$$

11.3.2 *The Fuchsian expansion*

We look for a solution of the original system which approaches the VTD solution as t tends to infinity. We write it in the following form, with the various $\varepsilon's$ being positive numbers to be chosen later:

$$\sigma^{ab} = \tilde{\sigma}^{ab} + e^{-\varepsilon_\sigma t}u^{ab}_\sigma \tag{11.33}$$

$$k^b_a = e^{-\lambda}\left(\frac{1}{2}\tilde{v}\delta^b_a + e^{-\varepsilon_k t}u^b_{k,a}\right); \quad \text{hence} \quad \tau := k^a_a \equiv e^{-\lambda}(\tilde{v} + e^{-\varepsilon_k t}u^a_{k,a}). \tag{11.34}$$

We take as a gauge condition

$$e^{-\varepsilon_k t}u^a_{k,a} + \frac{1}{2}\psi = 0 \tag{11.35}$$

Comparing the expressions (11.34) and (11.11) for τ shows that this condition is equivalent to the hypothesis that

$$\partial_t\lambda = -\tilde{v}, \quad \text{hence} \quad \lambda = \tilde{\lambda} - \tilde{v}t \tag{11.36}$$

with $\tilde{\lambda}$ independent of t but otherwise arbitrary.

We set

$$\phi = \hat{\phi} + e^{-\varepsilon_\phi t}u_\phi \equiv \tilde{\phi} - \frac{1}{2}\tilde{v}t + e^{-\varepsilon_\phi t}u_\phi. \tag{11.37}$$

We consider $\tilde{\lambda}$, $\tilde{\phi}$, $\tilde{v}$ as known functions on $\Sigma \times R$, independent of t. The Einstein scalar field equations are then a partial differential system for the unknowns u_σ, u_k, u_ϕ. In order to apply a Fuchsian theorem we write this system in first-order form. To this aim we introduce as usual new unknowns ϕ_t, ϕ_a, σ_c^{ab} satisfying the equations

$$\partial_t \phi = \phi_t, \quad \partial_t \phi_a = \partial_a \phi_t, \tag{11.38}$$

$$\partial_t \sigma_c^{\alpha b} = \tilde{\nabla}_c \partial_t \sigma^{ab} \tag{11.39}$$

with $\tilde{\nabla}_c$ the covariant derivative in the metric $\tilde{\sigma}$. We set

$$\phi_t \equiv -\frac{1}{2}\tilde{v} + e^{-\varepsilon_\phi t} u_{\phi_t}, \quad \phi_a \equiv \partial_a \tilde{\phi} - \frac{1}{2} t \partial_a \tilde{v} + e^{-\varepsilon_{\phi_a} t} u_{\phi_a} \tag{11.40}$$

$$\sigma_c^{ab} \equiv e^{-\varepsilon_{\sigma'} t} u^{ab}_{\sigma',c}. \tag{11.41}$$

The Einstein scalar evolution system is now a first-order system for the tensorial unknown $u := (u_\sigma, u_k, u_\phi, u_{\phi_t}, u_{\phi_a}, u_{\sigma'})$.

11.3.3 *First-order system*

The relation between g and k together with the Fuchsian expansion of k and σ give the equation[67]

$$\partial_t u_\sigma^{ab} - \varepsilon_\sigma u_\sigma^{ab} = 2 e^{(\varepsilon_\sigma - \varepsilon_k)t} \sigma^{ac} u^b_{k,c}. \tag{11.42}$$

The equation for σ_c^{ab} gives the following equation for u'_σ:

$$\partial_t u^{ab}_{\sigma',c} - \varepsilon_{\sigma'} u^{ab}_{\sigma',c} = e^{(\varepsilon_{\sigma'} - \varepsilon_\sigma)t} \tilde{\nabla}_c u^{ab}_\sigma, \tag{11.43}$$

of Fuchsian type as soon as $\varepsilon_{\sigma'} < \varepsilon_\sigma$.

After some manipulation (remark in particular that $e^{-\lambda}\tilde{v}^2$ disappears from the difference $\partial_t k_a^b - N\tau k_a^b$, as foreseen from the choice of the Fuchsian expansion), one finds that the Einstein evolution equation takes the form

$$\partial_t u^b_{k,a} - \varepsilon_k u^b_{k,a} - \frac{1}{2}\tilde{v}\delta^b_a u^c_{k,c} = e^{-\varepsilon_k t} u^c_{k,c} u^b_{k,a} + e^{\lambda + \varepsilon_k t} F^b_a(t, x, u, u_x), \tag{11.44}$$

where $F^b_a(t, x, u, D_x u)$ is a holomorphic function of x, the $u's$ and their space derivatives, which is at most a quadratic polynomial in t.

To obtain a system for the $u's$ in obvious Fuchsian form we split the above equation into its trace and its traceless part. It gives

$$\partial_t u^a_{k,a} - \varepsilon_k u^a_{k,a} - \tilde{v} u^a_{k,a} = e^{-\varepsilon_k t} u^c_{k,c} u^a_{k,a} + e^{\lambda + \varepsilon_k t} F^a_a(t, x, u, D_x u) \tag{11.45}$$

and for the traceless part ${}^T u^b_{k,a}$ we have

$$\partial_t \, {}^T u^b_{k,a} - \varepsilon_k \, {}^T u^b_{k,a} = e^{-\varepsilon_k t} u^c_{k,c} \, {}^T u^b_{k,a} + e^{\lambda + \varepsilon_k t} {}^T F^b_a(t, x, u, u_x). \tag{11.46}$$

[67] For details of computation see Choquet-Bruhat, Isenberg, and Moncrief (2005).

The evolution equation for u_{ϕ_t} is easily deduced from the wave equation satisfied by ϕ to be of the form

$$\partial_t u_{\phi_t} - \varepsilon_{\phi_t} u_{\phi_t} = e^{(\varepsilon_{\phi_t} - \varepsilon_k)t} F + e^{(\lambda + \varepsilon_{\phi_t} t)} F \tag{11.47}$$

where the $F's$ are holomorphic function of x, the $u's$ and their space derivatives, which are at most linear in t. The equations for u_ϕ and u_{ϕ_a} reduce to

$$\partial_t u_\phi = e^{-\varepsilon_{\phi_t} t} u_{\phi_t}, \quad \partial_t u_{\phi_a} - \varepsilon_{\phi_a} u_{\phi_a} = e^{(\varepsilon_{\phi_a} - \varepsilon_{\phi_t})t} \partial_a u_{\phi_t}. \tag{11.48}$$

11.3.4 *Fuchsian evolution*

We have obtained for u a system of the form

$$\partial_t u - Au = e^{-\mu t} f(t, x, u, D_x u) \tag{11.49}$$

where A is a diagonal operator, with positive eigenvalues if the $\varepsilon' s$, and $\tilde{v}$ are positive. The system is Fuchsian if $\mu > 0$, and f bounded.

Theorem 11.1 *If the VTD solution is such that $\partial_t \hat{\phi} > 0$, it is possible to choose the parameters ε in the Fuchsian expansions of g, k, ϕ and some of their derivatives so that the Einstein scalar evolution system written in first-order form for the unknown u is a Fuchsian system.*

Proof Recalling that $e^\lambda = e^{\tilde{\lambda} - \tilde{v}t}$, inspection of the equations show that the inequalities

$$\tilde{v} > \varepsilon_k > \varepsilon_\sigma > \varepsilon_{\sigma'}, \quad \varepsilon_k > \varepsilon_{\phi_t} > \varepsilon_{\phi_a} > 0 \tag{11.50}$$

imply the required condition on the right-hand side of the considered system with $\mu < Inf(\varepsilon_{\sigma'}, \varepsilon_{\phi_a})$. □

Corollary 11.2 *The Einstein scalar evolution system has an analytic solution tending as t tends to infinity to the VTD solution defined by the arbitrary analytic data $\tilde{\sigma}$, $\tilde{\lambda}$, $\tilde{\phi}$, $\tilde{v}$ on Σ.*

Proof The solution of the considered system satisfies the original evolution equations if $\phi_a = \partial_a \phi$ and $\sigma_c^{ab} = \tilde{\nabla}_c \sigma^{ab}$. The proof of this fact is obtained by remarking that the equations satisfied by ϕ_a and ϕ_t together with commutation of partial derivatives show that:

$$\partial_t(\phi_a - \partial_a \phi) = \partial_a \phi_t - \partial_a \partial_t \phi = 0; \tag{11.51}$$

hence $\phi_a - \partial_a \phi$ is independent of t. When t tends to ∞ it tends to zero because

$$\phi_a - \partial_a \phi = e^{-\varepsilon_{\phi_a} t} u_{\phi_a} - e^{-\varepsilon_\phi t}(\partial_a u_\phi - \varepsilon_\phi u_\phi). \tag{11.52}$$

It is therefore always zero. An analogous proof holds to show that $\sigma_c^{ab} = \tilde{\nabla}_c \sigma^{ab}$. □

11.3.5 *Constraints*

The solution of the evolution system satisfies the full Einstein equations if it satisfies the Einstein constraints, that is:

$$C_0 \equiv \Sigma_0^0 \equiv -\frac{1}{2}\{R(g) - k.k + \tau^2 - e^{-2\lambda}2|\partial_t\phi|^2\} = 0$$

$$C_a \equiv e^{\lambda}\Sigma_a^0 \equiv -\{\nabla_b k_a^b - \partial_a\tau + e^{-\lambda}2\partial_t\phi\partial_a\phi\} = 0.$$

As usual one uses the Bianchi identities to prove that a solution such that the VTD constraints are satisfied annul the full constraints during its time of existence.

11.4 *The unpolarized case*

11.4.1 *Wave map VTD solution*

The results for a VTD wave map are very different from the result obtained for a scalar function. If we drop space derivatives in the wave map equations we obtain geodesic equations in the target manifold, with t the length parameter on these geodesics when the 2+1 metric is in harmonic time. By the change of coordinates $Y = e^{2\phi}$ in the target, which is effectively a diffeomorphism from R^2 on to the upper half plane $Y > 0$, the metric G takes a standard form for the metric of a Poincaré half plane, namely

$$G \equiv \frac{1}{2}\left\{\frac{d\omega^2 + dY^2}{Y^2}\right\}, \quad Y = e^{2\phi}. \tag{11.53}$$

The VTD geodesic equations written in this metric read, with a prime denoting the derivative with respect to t:

$$\omega'' - 2Y^{-1}\omega'Y' = 0, \quad Y'' + Y^{-1}\omega'X' = 0. \tag{11.54}$$

The general solution of these geodesic equations is represented in these coordinates, as it is well known and can easily be checked, by half circles[68] centred on the line $Y = 0$, namely, with A and B arbitrary constants (that is independent of t):

$$\hat{\omega} = B + A\cos\theta, \quad \hat{Y} = A\sin\theta, \quad 0 < \theta < \pi. \tag{11.55}$$

These functions X and Y satisfy the differential equations (11.54) if and only if it holds that

$$\frac{\theta''}{\theta'} = \frac{\cos\theta}{\sin\theta}\theta'. \tag{11.56}$$

Integration gives

$$\theta' = -\tilde{w}\sin\theta \quad \text{and} \quad tg\frac{\theta}{2} = \tilde{\Theta}e^{-\tilde{w}t}. \tag{11.57}$$

[68] We discard here the special case which corresponds to the polarized case, treated above, where these circles are centred at infinity, the geodesics are then the half lines $\mathrm{X} \equiv \omega = \text{constant}$.

In the formulae (11.55) we denote $\frac{1}{2}logA$ by $\tilde{\phi}$ and B by $\tilde{\omega}$. They read then:

$$\hat{\phi} = \tilde{\phi} + \frac{1}{2}log(sin\theta), \qquad \hat{\omega} = \tilde{\omega} + e^{2\tilde{\phi}}cos\theta \ . \tag{11.58}$$

11.4.2 *Half polarization condition*

We remark that $\hat{Y} \equiv e^{2\hat{\phi}}$ tends to zero when t tends to ∞, but $\hat{\omega}$ tends to $\tilde{\omega}+e^{2\tilde{\phi}}$. We label as the **half polarization** condition[69] the equation satisfied on Σ by the VTD solution

$$\tilde{\omega} + e^{2\tilde{\phi}} = \text{constant}. \tag{11.59}$$

It is possible to construct AVTD solutions if and only if the VTD solution satisfies this non-generic condition[70]. Its geometric interpretation is that the geodesics corresponding to the VTD wave map are represented in the Poincaré plane by half circles with the same end point on the axis $Y = 0$.

[69] Generalization of a condition found by Isenberg and Moncrief (2002).
[70] See the proof in Choquet-Bruhat and Isenberg (2006).

XIV

STATIONARY SPACETIMES AND BLACK HOLES

1 Introduction

We call **stationary** an $n+1$ spacetime $(V, \hat{g})$ which admits a 1-parameter group $G_1 \equiv R$ of isometries with timelike [1] Killing vector ξ, whose orbits are diffeomorphic[2] to R and span the manifold V. We include in our definition of stationary the assumption that the quotient of V by the action of G_1 is a spacelike manifold examples where such a quotient do not exist are a torus[3] with an action of R in an irrotational direction or the case where the Killing vector vanishes at a point[4]. In mathematical terms $(V, \hat{g})$ is a principal fibre bundle with group and typical fibre R and base a Riemannian manifold (M, g). Since a principal bundle with fibre R^N is trivial, the manifold V is diffeomorphic to the product $M \times R$. If we identify V with one of its global trivializations, the metric $\hat{g}$ reads

$$\hat{g} \equiv -\psi^2 (dt + a)^2 + g \tag{1.1}$$

with $t \in R$ a time coordinate on the orbits such that $\xi = \frac{\partial}{\partial t}$; ψ, a, and g are respectively a scalar, a 1-form, and a Riemannian metric on the basis M. In local coordinates x^i in the domain U of a chart of M, one has

$$a \equiv a_i dx^i, \qquad g \equiv g_{ij} dx^i dx^j \tag{1.2}$$

with all coefficients in $\hat{g}$ being independent of t.

A stationary spacetime is called **static** if the vector spaces orthogonal to the orbits are globally integrable; i.e. the orbits are orthogonal to n-dimensional space manifolds. This is true if the 1-form a is an exact differential, $a = df$. The change of time parameter

$$t' = t + f \tag{1.3}$$

[1] Some authors call "strictly stationary" such spacetimes.

[2] We exclude the case of orbits diffeomorphic to S^1, i.e. the case of closed timelike curves, because it is considered as non-physical.

[3] Chruściel, P., private communication.

[4] T. Damour, private communication, see remark 1.2.

(i.e. choosing the origin for t on each orbit in one of these orthogonal submanifolds) puts the metric of a static spacetime in the form

$$\hat{g} \equiv -\psi^2 dt'^2 + g_{ij} dx^i dx^j, \tag{1.4}$$

with g and ψ independent of t'.

Static spacetimes are considered as representing equilibrium situations, while stationary spacetimes model permanent motions. Both play an important role in relativistic dynamics.

Remark 1.1 *Static spacetimes are invariant under time reversal, $t \to -t$. This property can also be taken as a definition to distinguish static spacetimes among stationary ones. One can also define spacetimes which are only locally static; i.e. the vector spaces orthogonal to the orbits are only locally integrable.*

Remark 1.2 A more general definition of stationarity permits the existence of physically interesting gravitational solitous, particularly in string theory[5]

In this chapter we give the proofs of fundamental uniqueness theorems for complete stationary solutions of the vacuum, or electrovac, solutions of the Einstein equations. We study the properties of the Kerr stationary black hole, of which the Schwarzschild black hole (see Section 4) is a particular case. We survey briefly the history of the research on the uniqueness theorem for 3+1-dimensional stationary black holes which was pictured by J. A. Wheeler using the picturesque phrase "black holes have no hair". The proof of the uniqueness of the Kerr solution as a stationary black hole has finally been completed[6], with a technical hypothesis on the horizon, but without analyticity hypothesis. The Einstein equations with stationary metric and sources are essentially elliptic, like the non-stationary Einstein equations with sources which satisfy hyperbolic equations are essentially hyperbolic in a sense we have explained before. It is well known that solutions of elliptic equations with analytic coefficients are analytic. Stationary vacuum Einsteinian spacetimes are analytic in well-chosen gauges. However, since the Killing vector of a stationary black hole turns null on the horizon, it is not legitimate to assume analyticity.

In this chapter, apart from analyticity, we present most of the theorems without being too precise about the smoothness assumptions they require.

The sections "Rigidity theorem" and "Further results" have been written respectively by James Isenberg and Piotr Chruściel.

2 Spacetimes with 1-parameter isometry group

We apply to the case where the isometry group of the metric is one-dimensional the general Kaluza–Klein reductions which are discussed in an appendix[7]. The

[5] See e.g. CB-DM1 IV.
[6] Ionescu, A. and Klainerman, S. gr;qc 0711.0040v1.
[7] See also the Section IX.14.

formulae can also be computed[8] directly; we leave that as an exercise to the reader. One-parameter timelike isometry groups define stationary spacetimes, but one-parameter spacelike isometry group are relevant in 3+1 General Relativity, and also in 4+1-dimensional spacetimes, leading then to a unification of gravity with electromagnetism, which is the original Kaluza–Klein reduction. We treat in this section the general case of a pseudo-Riemannian manifold, in view of later applications.

We consider a spacetime $(V, \hat{g})$ with V a smooth $n+1$-dimensional manifold. We suppose that V is a principal fibre bundle with base an n-dimensional manifold M and group G_1 a 1-parameter connected Lie group, i.e. R or S^1. We suppose that the metric $\hat{g}$ is invariant under the right action of G_1. The corresponding Killing vectors, tangent to the orbits of the group are supposed to be all timelike or all spacelike. We denote by π the bundle projection $V \to M$, by π^* the pull back from tensors on M into tensors on V. To treat simultaneously the cases where the Killing field e_m is timelike or spacelike we write the spacetime metric g as follows

$$\hat{g} \equiv \pi^* g + (\pi^* \Phi)\theta^2, \tag{2.1}$$

where g is a pseudo-Riemannian metric and Φ a function, both on M[9], and θ is a G_1 connection 1-form on V.

We consider a local trivialization $U \times G_1$ of V, $U \subset M$ with local coordinates x^a, and we denote by x^m a coordinate in G_1. In such coordinates $\pi^* g$ and g have the same expression

$$\pi^* g = g = g_{ab} dx^a dx^b \tag{2.2}$$

while θ is the 1-form

$$\theta = dx^m + A_a dx^a, \tag{2.3}$$

the Killing vector field ξ being

$$\xi \equiv e_m \equiv \frac{\partial}{\partial x^m}. \tag{2.4}$$

The functions g_{ab}, Φ, A_a are all independent of x^m. Note that $|\Phi|^{\frac{1}{2}}$ is the length of the Killing field ξ and that $A := A_a dx^a$ is a 1-form on U representing the G_1

[8] The structure constants of the group are all zero, and the latin indices $m, n, \ldots$ reduce to one index $n+2$.

[9] g is Riemannian and Φ negative for stationary spacetimes; g is Lorentzian and Φ positive if the isometry group is spacelike.

connection in the trivialization $U \times G_1$; this 1-form A changes under a change of trivialization, $R \to R$ by $x^m \mapsto h(x^a)x^m$, via the formula

$$A \to A + h^{-1}dh. \tag{2.5}$$

The 1-form A can be defined on the whole of M if and only if the bundle $V \to M$ is trivial (i.e. $V = M \times G_1$). This is always the case when $G_1 = R$, but not when[10] $G_1 = S^1$.

The curvature of the G_1 connection is the 2-form on M

$$f = dA, \quad \text{i.e.} \quad \text{in coordinates} \quad f_{ab} = \nabla_a A_b - \nabla_b A_a. \tag{2.6}$$

The vanishing of the differential df of a 2-form f on M is a necessary condition for f to be the differential of a 1-form A. It is not sufficient when M is not diffeomorphic to R^n (see e.g. CB-DM1 IV B, and also Section XVI.3).

2.1 *Connection and Riemann tensor*

In the case of a 1-parameter Lie group the general Kaluza–Klein formulas simplify. The only non-zero structure coefficients of the coframe (dx^a, θ) are

$$\hat{c}^m_{ab} = -f_{ab} \tag{2.7}$$

The connection coefficients are computed to be

$$\hat{\omega}^a_{bc} = \Gamma^a_{bc}, \tag{2.8}$$

where Γ^a_{bc} are the Christoffel symbols of the metric g on M,

$$\hat{\omega}^m_{mm} = 0, \quad \hat{\omega}^m_{am} = \frac{1}{2}\Phi^{-1}\partial_a\Phi, \quad \hat{\omega}^a_{mm} = -\frac{1}{2}g^{ab}\partial_b\Phi. \tag{2.9}$$

$$\hat{\omega}^m_{ab} = -\frac{1}{2}f^m_{ab}, \quad \hat{\omega}^a_{mb} = \hat{\omega}^a_{bm} = -\frac{1}{2}f^a{}_{b,m}, \tag{2.10}$$

Here the indices a, b are raised with g^{ab}, and the index m is lowered with Φ; we have set

$$f^m_{ab} := f_{ab}, \quad \text{hence} \quad f_{ab,m} := \Phi f_{ab}. \tag{2.11}$$

2.2 *Curvature tensor*

We find by computation, with $R_{cd}{}^a{}_b$ the curvature tensor of g,

$$\hat{R}_{cd}{}^a{}_b \equiv R_{cd}{}^a{}_b + \frac{1}{2}f_{m,b}{}^a f^m_{cd} + \frac{1}{4}f_{m,c}{}^a f^m_{bd} - \frac{1}{4}f_{m,d}{}^a f^m_{bc} \tag{2.12}$$

$$\hat{R}_{cm}{}^a{}_m \equiv \nabla_c\hat{\omega}^a_{mm} - \hat{\omega}^a_{md}\hat{\omega}^d_{cm}; \tag{2.13}$$

[10] The Einstein–Maxwell equations are obtained by a Kaluza–Klein reduction of five-dimensional Einstein equations with group S^1. A non-trivial bundle invalidates the non-existence theorem of solitons.

with ∇ the covariant derivative in the metric g. Hence

$$\hat{R}_{cm}{}^{a}{}_{m} \equiv -\frac{1}{2}\nabla_c\nabla^a\Phi + \frac{1}{4}f_{m,d}{}^{a}f_{m,c}{}^{d} - \frac{1}{4}\Phi^{-1}\partial^a\Phi D_c\Phi. \tag{2.14}$$

On the other hand, with $[AB] = AB - BA$

$$\hat{R}_{cd}{}^{a}{}_{m} \equiv \frac{1}{2}\nabla_{[c}f_{m,d]}{}^{a} + \frac{1}{2}g^{ab}\partial_b\Phi f^{m}_{cd} + \frac{1}{4}f_{m,[c}{}^{a}\Phi^{-1}\partial_{d]}\Phi. \tag{2.15}$$

2.3 *Ricci tensor*

Straightforward computation gives:

$$\hat{R}_{ab} \equiv R_{ab} - \frac{1}{2}f_{m,b}{}^{c}f^{m}_{ac} - \frac{1}{2}\Phi^{-1}\nabla_a\partial_b\Phi + \frac{1}{4}\Phi^{-2}\partial_a\Phi\partial_b\Phi, \tag{2.16}$$

$$\hat{R}_{am} \equiv \frac{1}{2}\nabla_b f_{m,a}{}^{b} + \frac{1}{4}f_{m,a}{}^{b}\Phi^{-1}\partial_b\Phi, \tag{2.17}$$

$$\hat{R}_{mm} \equiv \frac{1}{4}\{f_{m,ab}f^{ab}_{m} + \Phi^{-1}\partial_a\Phi\partial^a\Phi - 2\nabla_a\nabla^a\Phi\} \tag{2.18}$$

If $\Phi > 0$; i.e. if the isometry group is spacelike, these formulae can be rewritten as follows

$$\hat{R}_{am} \equiv \frac{1}{2\Phi^{\frac{1}{2}}}\nabla_b(f_{m,a}{}^{b}\Phi^{\frac{1}{2}}) \equiv \frac{1}{2}\nabla_b(f_a{}^{b}\psi^{\frac{3}{2}}), \quad \psi := \Phi^{\frac{1}{2}} > 0,. \tag{2.19}$$

$$\hat{R}_{mm} \equiv \frac{1}{4}\psi^4 f_a{}^{b}f_b{}^{a} - \psi\nabla^a\partial_a\psi. \tag{2.20}$$

3 Stationary spacetimes

3.1 *General case*

3.1.1 *Equations*

If the Killing vector ξ of the isometry group is timelike we return to our usual notations $x^m = x^0 = t$ for the coordinate on the group orbit, and x^i for coordinates on the spacelike base space M. We set $\Phi := \xi^\alpha\xi_\alpha \equiv -\psi^2$, the spacetime metric reads

$$\hat{g} := -\psi^2(dt + a_i dx^i)^2 + g_{ij}dx^i dx^j. \tag{3.1}$$

The formulas of the previous section, or direct computation[11] of the components of the Ricci tensor of the above spacetime metric, show that the Einstein

[11] Lichnerowicz, A. (1955) *Théories Relativistes de la Gravitation et de l'éléctromagnétisme*, Masson, computes analogous formulae in an orthonormal frame.

equations with source ρ reduce to

$$\hat{R}_{ij} \equiv R_{ij} + \frac{\psi^2}{2} f_i{}^h f_{jh} - \psi^{-1}\nabla_i\partial_j\psi = \rho_{ij}, \tag{3.2}$$

$$\hat{R}_{i0} \equiv -\frac{1}{2\psi}\nabla_j(f_i{}^j\psi^3) = \rho_{i0}, \tag{3.3}$$

$$\hat{R}_{00} \equiv \frac{1}{4}\psi^4 f_i{}^j f_j{}^i + \psi\Delta_g\psi = \rho_{00}. \tag{3.4}$$

3.1.2 *Divergence formulae*

The following identities are important for the proof of some properties of stationary spacetimes. The first one is a trivial consequence of Equations (3.3). The second one[12] is a generalization of the **Gauss formula** of classical mechanics.

Theorem 3.1 *In a stationary spacetime* $(M \times R, \hat{g})$

1. *The following identity holds*

$$\psi a^i \hat{R}_{i0} \equiv -\frac{1}{2}\nabla_j(a^i f_i{}^j\psi^3) + \frac{1}{4}\psi^3 f_i{}^j f_j{}^i. \tag{3.5}$$

2. *The time–time component* R^t_t *of the Ricci tensor of a stationary spacetime* $(M \times R, \hat{g})$ *in a natural frame with time axis tangent to the timelines and space axis tangent to the manifold* M_t *satisfies the identity*

$$R^t_t \equiv R^t_t \equiv -\psi^{-1}\nabla^i\left\{\partial_i\psi + \frac{1}{2}(a^j f_{ji}\psi^3)\right\} \tag{3.6}$$

Proof

1. We use the elementary result

$$a^i\nabla_j(f_i{}^j\psi^{3\cdot}) \equiv \nabla_j(a^i f_i{}^j\psi^3) - f_i{}^j\psi^3\nabla_j a^i \tag{3.7}$$

   and the definition of $f_i{}^j$.

2. The considered natural frame is the dual of the coframe (dt, dx^i) while the coframe used in Equations (3.1–4) is

$$\theta^0 = dt + a_i dx^i, \quad \theta^i = dx^i. \tag{3.8}$$

   The change of frame formula gives

$$R^t_t \equiv \hat{R}^0_0 + a_i\hat{R}^i_0. \tag{3.9}$$

   The identities (3.2–4) imply that

$$\hat{R}^0_0 \equiv -\psi^{-2}\hat{R}_{00}, \quad \hat{R}^i_0 \equiv g^{ij}\hat{R}_{0j} \tag{3.10}$$

[12] Fourès (Choquet)-Bruhat, Y. (1948) *C. R. Acad. Sci. Paris*, **226**, 48–51.

Therefore, by (3.9),

$$R^t_t \equiv -\left\{\frac{1}{4}\psi^2 f_i{}^j f_j{}^i + \psi^{-1}\Delta_g\psi + \frac{1}{2}\psi^{-1}a^i\nabla_j(f_i{}^j\psi^3)\right\} \tag{3.11}$$

which gives, using (3.7),

$$R^t_t \equiv -\psi^{-1}\{\Delta_g\psi + \frac{1}{2}\nabla_j(a^i f_i{}^j\psi^{3\cdot})\} \tag{3.12}$$

that is, the identity (3.6). □

3.2 *Static spacetimes*

In a static spacetime with the metric written in the canonical form

$$\hat{g} = -\psi^2 dt^2 + g_{ij}dx^i dx^j, \tag{3.13}$$

the time adapted frame coincides with a Cauchy adapted frame with lapse $N = \psi$ and with zero shift. The extrinsic curvature of a static spacetime is $K = 0$, and the Einstein equations reduce to

$$\hat{R}_{ij} \equiv R_{ij} - \psi^{-1}\nabla_i\partial_j\psi = \rho_{ij}, \tag{3.14}$$

$$\hat{R}_{i0} \equiv 0 = \rho_{0i}, \tag{3.15}$$

$$\psi^{-1}\hat{R}_{00} \equiv \Delta_g\psi = \psi^{-1}\rho_{00}. \tag{3.16}$$

The second equation shows that the source must have a zero momentum for the spacetime to be static, justifying its name. A reciprocal theorem is easy to prove under physically meaningful hypothesis.

To be short we label as **stationary asymptotically Euclidean spacetimes** those stationary spacetimes $(V, \hat{g})$ such that the space n-manifold M is the union of a compact set with the outside Ω of a ball of Euclidean space (R^n, e); $g - e$, a, and $1 - \psi$ fall off at infinity like r^{-p} while their derivatives of order k fall off like $r^{-(p+k)}$, with[13] $p > \frac{n-2}{2}$.

Theorem 3.2 *A stationary spacetime with sources of zero momentum, i.e.* $\rho_{0i} \equiv 0$, *is locally static if either*

1. *M is compact.*
2. *The spacetime is asymptotically Euclidean.*

Proof If $\rho_{0i} = 0$, the identity (3.5) implies

$$\nabla_j(a^i f_i{}^j\psi^{3\cdot}) = \frac{1}{2}\psi^3 f_i{}^j f_j{}^i. \tag{3.17}$$

[13] The definition is meaningful only for $n > 2$.

If M is compact, the integration of the above equality on M gives the result.

If M is non-compact, the integration on a ball $B_\rho \subset M$ and the Stokes formula gives

$$\int_{\partial B_\rho} a^i f_i{}^j \psi^3 n_j \mu_{\partial B_\rho} = \frac{1}{2} \int_{B_\rho} \psi^3 f_i{}^j f_j{}^i \mu_{B_\rho}. \tag{3.18}$$

When the radius ρ of the ball tends to infinity the left-hand side tends to zero if $n > 2$, because if the spacetime is asymptotically Euclidean, the product af falls off like $r^{-(2p+1)}$, $2p + 1 > n - 1$, while $\mu_{\partial B_\rho} \tilde{=} r^{n-1} \mu_S$ where μ_S is the volume element of the sphere S^{n-1}. □

4 Gravitational solitons

One labels a gravitational **soliton** a complete non-trivial (i.e. non-flat) stationary solution of the vacuum Einstein equations.

4.1 *Elementary proofs*

The following theorem is easy to prove.

Theorem 4.1 *The only static solutions of the Einstein equations with compact spacelike sections and source satisfying the strong energy condition are vacuum solutions.*

Proof Integration of (3.16) on the compact space manifold M gives

$$\int_M \psi^{-1} \rho_{00} \mu_g = 0. \tag{4.1}$$

Therefore $\rho_{00} \equiv 0$ if $\psi > 0$ and $\rho_{00} \geq 0$ on M. □

We extend as follows to arbitrary dimensions an elementary proof of a theorem given in $3 + 1$ dimensions by Einstein and Pauli in the static case and by Lichnerowicz in the stationary case.

Theorem 4.2 *In $n+1$-dimensional spacetimes the gravitational solitons $(M \times R, \hat{g})$ with M compact are locally static with Ricci flat space metric and ψ a constant.*

The same result holds when it is only supposed that (M, g) is complete with the additional hypothesis that ψ attains its minimum at an interior point of M.

Proof In a vacuum spacetime Equations (3.3), (3.4), (3.2) read

$$\nabla_j (f_i{}^j \psi^{3\cdot}) = 0 \tag{4.2}$$

$$\Delta_g \psi = -\frac{1}{4} \psi^3 f_i{}^j f_j{}^i, \tag{4.3}$$

$$R_{ij} = \psi^{-1} \nabla_i \partial_j \psi - \frac{\psi^2}{2} f_i{}^h f_{jh}. \tag{4.4}$$

1. Compact M. The integration of (4.2) on M implies $f = 0$, hence V is locally static. The equation $\Delta_g \psi = 0$ on a compact manifold M implies $\psi =$ constant by the maximum principle.

If $\psi =$ constant, Equations (4.4) imply $Ricci(g) = 0$.

2. (M, g) is a complete Riemannian manifold. The maximum principle implies that ψ cannot attain a minimum at an interior point without being a constant. If ψ is constant then (4.3) implies that $f_i^j = 0$, i.e. the spacetime is locally static, hence $Ricci(g) = 0$ by (4.4). □

Corollary 4.3 *If $(V, \hat{g})$ is asymptotically Euclidean, the hypothesis on ψ attaining its minimum at an interior point is unnecessary.*

Proof We integrate, for a vacuum spacetime, the divergence identity (3.5) in a ball of M large enough to have its boundary ∂B_ρ in the domain $R^n - K$ where the spacetime metric is asymptotically Euclidean. The Stokes formula gives

$$\int_{\partial B_\rho} a^i f_i{}^j \psi^3 \cdot n_j \mu_{\partial B_\rho} = \int_{B_\rho} \frac{1}{2} \psi^3 f_i{}^j f_j{}^i \mu_\rho \tag{4.5}$$

The left-hand side tends to zero as r tends to infinity; the same property holds therefore for the right-hand side. It implies $f_i{}^j f_j{}^i = 0$, hence the spacetime is locally static. □

These theorems are independent of the spacetime dimension. The following one is restricted to four-dimensional spacetimes.

Corollary 4.4 *Under the hypotheses of Theorem 4.2 or of its Corollary (4.3), the only $3+1$-dimensional gravitational solitons are flat. They are local direct metric products of R, with metric $-dt^2$, with a flat Riemannian 3-manifold. The only gravitational soliton is Minkowski space M^4 if the space manifold is R^3.*

Proof In three dimensions a Ricci flat manifold is flat. □

Remark 4.5 *There are, besides the flat torus T^3, five non-isometric (i.e. non-diffeomorphic) Riemannian orientable compact flat 3-manifolds, quotients of the Euclidean space E^3 by discrete isometry groups*[14].

M. T. Anderson[15] gives, for four-dimensional spacetimes, a more general theorem. His study involves in particular the Cheeger–Gromov theory of the collapse of Riemannian manifolds, and is outside the scope of this book.

[14] See Wolf, J. *Spaces of Constant Curvature*, Chapter 3, Publish or Perish Inc., Houston TX.

[15] Anderson, M. T. (2001) *Commun. Math. Phys.*, **222**(3), 533–67, and references therein.

4.2 *Case $n = 3$, Komar mass*

The **Komar mass** is defined for $3+1$-dimensional stationary asymptotically Euclidean spacetimes. It has a geometric formulation and useful properties for the study of black holes.

The Komar mass agrees with the mass which appears in the coefficient g_{00} which plays the role of the potential in the Newtonian approximation of General Relativity (see Chapter 3). For stationary spacetimes the Komar mass agrees with the ADM mass.

4.2.1 *Definitions*

Let $(V, \hat{g})$ be a stationary $3+1$-dimensional spacetime with base manifold M and Killing vector ξ. The metric adjoint $*d\xi$ of the 2-form $d\xi$,

$$d\xi \equiv \frac{1}{2}(\hat{\nabla}^\lambda \xi^\mu - \hat{\nabla}^\mu \xi^\lambda) dx^\lambda \wedge dx^\mu, \tag{4.6}$$

is also a 2-form, given by[16] (we use the Killing property $\hat{\nabla}^\lambda \xi^\mu + \hat{\nabla}^\mu \xi^\lambda = 0$)

$$*d\xi := \hat{\eta}_{\alpha\beta\lambda\mu} \hat{\nabla}^\lambda \xi^\mu dx^\alpha \wedge dx^\beta. \tag{4.7}$$

We have defined (Appendix I) an asymptotically Euclidean manifold to be a complete Riemannian manifold, the union of a compact set and the exterior of a ball[17] of R^n with an asymptotically Euclidean metric. We now consider also non-complete spaces (M, g), where the complement of the ball of R^n is non-compact and we call them **asymptotically Euclidean manifolds with non-compact interior**.

If the manifold M can be extended to a manifold with compact boundary ∂M and g extends at least continuously to ∂M, we call $(M \cup \partial M, g)$ an **asymptotically Euclidean manifold with boundary**.

Definition 4.6 *Let (M, g) be a Riemannian asymptotically Euclidean manifold, possibly with non-compact interior, which is the base of a stationary spacetime $(V, \hat{g})$ with Killing timelike vector ξ. The **Komar mass** of $(V, \hat{g})$ is the integral*

$$m_{Komar} := lim_{r\to\infty} \frac{1}{4\pi} \int \int_{S^2 \times \{r\}} *d\xi. \tag{4.8}$$

The following lemma is fundamental for application to Einsteinian spacetimes.

Lemma 4.7 *In a stationary $3+1$-spacetime with Killing vector ξ, the differential of the 2-form $*d\xi$ is the 3-form adjoint of the 1-form $4Ricci(g).\xi$.*

Proof The differential of the 2-form $*d\xi$ is the 3-form

$$d * d\xi \equiv \hat{\eta}_{\alpha\beta\lambda\mu} \hat{\nabla}_\gamma \hat{\nabla}^\lambda \xi^\mu dx^\gamma \wedge dx^\alpha \wedge dx^\beta \tag{4.9}$$

[16] See Heusler, M. (1996) *Black Hole Uniqueness Theorems*, Cambridge University Press.

[17] To avoid complications in writing we consider asymptotically Euclidean spaces with only one end.

Its adjoint is the 1-form

$$*d*d\xi \equiv \hat{\eta}_{\rho\gamma\alpha\beta}\hat{\eta}^{\alpha\beta\lambda\mu}\hat{\nabla}^{\gamma}\hat{\nabla}_{\lambda}\xi_{\mu}dx^{\rho} = 4(\hat{\nabla}^{\gamma}\hat{\nabla}_{\rho}\xi_{\gamma} - \hat{\nabla}_{\rho}\hat{\nabla}^{\gamma}\xi_{\gamma})dx^{\rho} \tag{4.10}$$

$$= 4\hat{R}_{\rho\mu}\xi^{\mu}dx^{\rho}. \tag{4.11}$$

□

Corollary 4.8 *Let M be a space slice with equation $t =$ constant in a trivialization $V = M \times R$ of a stationary spacetime $(V, \hat{g})$ with Killing vector ξ. The integral on a vacuum domain of M of the 3-form $d*d\xi$ is zero.*

Proof The proof is trivial since the lemma shows that $Ricci(\hat{g}) = 0$ in a domain implies $d*d\xi = 0$ in this domain. □

The following lemma relates the Komar mass of an Einsteinian spacetime with an integral on the boundary of a stationary asymptotically Euclidean spacetime with boundary.

Lemma 4.9 *Let $(V, \hat{g})$ be a 3+1 vacuum Einsteinian stationary spacetime with Killing vector ξ and with basis an asymptotically Euclidean manifold M with boundary ∂M. Then, with the proper orientation of the boundary, the Komar mass of $(V, \hat{g})$ is given by*

$$m_{Komar} = \frac{1}{4\pi}\int\int_{\partial M} *d\xi. \tag{4.12}$$

Proof Use the Stokes formula and the above Corollary 4.8. □

The formula (4.12) is taken as the definition of the mass of a stationary black hole in a vacuum spacetime.

4.2.2 *Computation and positivity*

To compute the Komar mass, we use adapted natural coordinates on V; that is, $\frac{\partial}{\partial t} = \xi$, and e_i natural frame in M. Then $\xi^0 = 1$, $\xi^i = 0$. The metric $\hat{g}$ reads

$$\hat{g} = -\psi^2 dt^2 - 2\psi a_i dx^i dt + (g_{ij} - a_i a_j)dx^i dx^J. \tag{4.13}$$

The components of the covariant vector (1-form) associated with ξ, denoted also by ξ, are

$$\xi_0 = -\psi^2, \tag{4.14}$$

$$\xi_i = -a_i\psi. \tag{4.15}$$

Hence

$$d\xi = 2\psi\partial_i\psi dx^i \wedge dt - \frac{1}{2}\psi f_{ij}dx^i \wedge dx^j \tag{4.16}$$

The adjoint of the 2-form $-\frac{1}{2}\psi f_{ij}dx^i \wedge dx^j$ on a $3+1$-dimensional manifold contains a factor dt, and hence vanishes on the manifold M, $t =$ constant.

To estimate the adjoint of $d\xi$ we assume[18] that, asymptotically, with m some constant[19],

$$\psi \doteq 1 - \frac{m}{r} + o\left(\frac{1}{r}\right), \quad \text{and} \quad \partial_i \psi = \frac{m}{r^2}\frac{x^i}{r} + o\left(\frac{1}{r^2}\right). \tag{4.17}$$

We remark that $-\frac{m}{r}$ is the asymptotic value of the Newtonian potential of a compactly supported body of mass m in Euclidean space E^3 and we call this m the asymptotic Newtonian mass of the stationary spacetime (the m appearing in the $g_{00} \equiv 1 - \frac{2m}{r} \doteq (1 - \frac{m}{r})^2$ coefficient of the Schwarzschild metric is its asymptotic Newtonian mass).

The restriction of $*d\xi$ to an $r =$ constant, two-dimensional submanifold of M, is asymptotically

$$*d\xi|_r \doteq \frac{\partial \psi}{\partial r}\omega_{S^2} \tag{4.18}$$

with ω_{S^2} the volume form of a sphere S^2 of radius r. We have proved, under the stated assumptions, the following lemma.

Lemma 4.10 *The Komar mass of an asymptotically Euclidean spacetime is equal to its asymptotically Newtonian mass.*

We can now prove the following theorem.

Theorem 4.11 *Let $(V = M \times R, \hat{g})$ be a $3+1$-dimensional asymptotically Euclidean[20] Einsteinian stationary spacetime. Suppose the Killing vector ξ such that*

$$Ricci(\hat{g})(\xi, \xi) \geq 0 \quad \textit{and is integrable on} \quad M. \tag{4.19}$$

Then the Komar mass of $(V, \hat{g})$, hence also the coefficient m in ψ, is non-negative. The Komar mass, hence also m, is equal to zero if and only if the spacetime is vacuum.

Proof When M is complete the Stokes formula implies that

$$m_{Komar} = \frac{1}{4\pi}\int_M d * d\xi. \tag{4.20}$$

We have seen that $d * d\xi$ is the adjoint of the 1-form $4\hat{R}_{\rho\mu}\xi^\mu dx^\rho$. The restriction to the manifold M of this 3-form is

$$d * d\xi_{|M} = 4\hat{R}_{00}\omega_M \equiv 4\hat{R}icci(\xi, \xi)\omega_M \tag{4.21}$$

[18] Such asymptotic expansions are argued to be justified in stationary spacetimes by Beig, R. and Simon, W. (1981) *Proc. R. Soc. Lond. Ser. A*, **376**, 333–41, following earlier work of Geroch. See also a paper by Beig, R. and Chruściel, P. (2006) gr-qc 0612012v2, and the comments in Chapter 13 after Definition (XIII.6.1) of spaces harmonically flat at infinity.

[19] which we want to show to be positive!

[20] Hence complete in our terminology.

with ω_M the volume form of M. The 3-form $d*d\xi$ has therefore a positive integral on M if the sources satisfy the strong energy condition, since then $\hat{R}(\xi,\xi) \geq 0$. The non-negativity of m and its vanishing in vacuum results from the equalities (4.19) and (4.20). □

The vanishing of m in vacuum and the fact that then $\partial_i\psi$ falls off faster than r^{-2} gives another proof[21] of the local staticity of stationary asymptotically Euclidean four-dimensional spacetimes, because the Stokes formula gives

$$\int_M \Delta_g \psi \omega_M = lim_{r\to\infty} \int\int_{S^2\times\{r\}} n^i \partial_i \psi r^2 \omega_{S^2} = 0, \tag{4.22}$$

from which follows, using (4.3), $f_i^j = 0$

4.2.3 *Komar mass and ADM mass*

The ADM mass of an asymptotically Euclidean spacetime is given by the formula (see "Positive mass theorems" published in DeWitt, C. and DeWitt, B. (eds.) (1983) *Relativity Groups and Topology*, Gordon and Breach),

$$m_{ADM}(g) := \frac{1}{16\pi} \lim_{r=\infty} \int_{S^2\times\{r\}} (\partial^j g_{ij} - \partial_i g^{jj}) n_i r^2 \omega_{S^2}. \tag{4.23}$$

Before proving the equality of the Komar mass of a *stationary* asymptotically Euclidean four-dimensional Einsteinian spacetime with its ADM mass[22], we prove the following lemma[23]. We call the ADM mass of any three-dimensional asymptotically Euclidean manifold with possibly non-empty interior, the number defined by the ADM formula[24].

Lemma 4.12 *The ADM mass of a three-dimensional asymptotically Euclidean manifold with metric* γ, $\gamma_{ij} = \delta_{ij} + O(\frac{1}{r})$, $\partial_h \gamma_{ij} = O(\frac{1}{r^2})$ *is equal to the following integral over a sphere at infinity*

$$m_{ADM}(\gamma) = \frac{1}{16\pi} lim_{r\to\infty} \int_{S^2\times\{r\}} S_{ij} n^i x^j r^2 \omega_{S^2}, \tag{4.24}$$

where S_{ij} *are the components of the Einstein tensor of* γ.

Proof The Einstein tensor is the sum of terms linear in the second derivatives of the metric and terms quadratic in the first derivatives, t_{ij}, called (up to sign)

[21] Though with stronger hypotheses and only for four-dimensional spacetimes.

[22] See another more direct proof in Chruściel, P. (1986) *Class. Quant. Grav.*, L115–21.

[23] Beig, R. (1978) *Phys. Rev. Lett. A*, **69**(3), 153–5.

[24] The ADM mass m_{ADM} of an asymptotically Euclidean manifold is positive or zero if the scalar curvature of this manifold is positive or zero, $\mathrm{m}_{ADM} = 0$ implies that the manifold is flat (see the "Positive mass theorems" article referred to above). These conclusions do not hold without the hypothesis on the scalar curvature.

the **Landau–Lifshitz pseudo tensor;** under our asymptotic hypotheses one has $t_{ij} = O(r^{-4})$, hence

$$lim_{r\to\infty} \int_{S^2\times\{r\}} t_{ij} n^i x^j r^2 \omega_{S^2} = 0. \tag{4.25}$$

Straightforward calculation shows that the quasilinear terms are divergence of a "superpotential", namely

$$S_{ij} + t_{ij} \equiv (det\gamma)^{-1} \partial_k h^{ijk}, \tag{4.26}$$

with

$$h^{ijk} := \partial_h [det\gamma(\gamma^{ij}\gamma^{kl} - \gamma^{jk}\gamma^{il})]. \tag{4.27}$$

Antisymmetry of h in i and j implies that

$$\partial_k \partial_j h^{ijk} \equiv 0 \quad \text{hence} \quad \int_{R^3} x^i \partial_k \partial_j h^{ijk} dx^1 dx^2 dx^3 \equiv 0. \tag{4.28}$$

Partially integrating this identity gives

$$lim_{r\to\infty} \int_{S^2\times\{r\}} S_{ij} n^i x^j r^2 \omega_{S^2} - \int_{R^3} \partial_k h^{iik} dx^1 dx^2 dx^3 = 0. \tag{4.29}$$

The last term is equal to a surface integral at infinity, found to be equal to

$$\lim_{r=\infty} \int_{S^2\times\{r\}} (\partial^j \gamma_{ij} - \partial_i \gamma^{jj}) n_i r^2 \omega_{S^2} \tag{4.30}$$

by straightforward computation using the definition (4.22). □

Theorem 4.13 *The Komar mass of a stationary asymptotically Euclidean Einsteinian spacetime with sources of compact support is equal to its ADM mass.*

Proof One relates the norm ψ of the timelike Killing vector ξ with the space metric (via the Einstein equations) by introducing a conformal space metric $\tilde{g}$, namely one sets

$$g = \psi^{-2}\tilde{g} \tag{4.31}$$

Then

$$\partial_i g_{hk} \equiv \psi^{-2}(\partial_i \tilde{g}_{hk} - 2\tilde{g}_{hk}\psi^{-1}\partial_i\psi) \tag{4.32}$$

Therefore, using the definition of the ADM mass and the hypothesis $\psi = 1 + 0(\frac{1}{r})$, $\partial_i\psi = 0(\frac{1}{r^2})$, $g_{ij} = \delta_{ij} + 0(\frac{1}{r})$

$$m_{ADM}(g) = m_{ADM}(\tilde{g}) - \lim_{r=\infty} \int_{S^2\times\{r\}} (\partial_i\psi - 3\partial_i\psi) n^i r^2 \omega_{S^2} = \tag{4.33}$$

$$m_{ADM}(g) = m_{ADM}(\tilde{g}) + 2 \lim_{r=\infty} \int_{S^2\times\{r\}} \partial_i\psi n^i r^2 \omega_{S^2}. \tag{4.34}$$

That is by the definition of the Komar mass

$$m(g)_{ADM} = m_{ADM}(\tilde{g}) + m_{Komar}. \tag{4.35}$$

Therefore $m_{ADM} = m_{Komar}$ if and only if $m_{ADM}(\tilde{g}) = 0$.

One deduces from the relations between the Ricci tensors of the conformal metrics g and $\tilde{g} = \psi^2 g$ (see Appendix VI) and the equations giving $Ricci(g)$ in a stationary spacetime that $\tilde{S}_{ij}$ tends to zero as r^{-4} when r tends to infinity if it is so of the sources. The formula (4.23) shows therefore that $m_{ADM}(\tilde{g}) = 0$, which completes the proof. □

Exercise. *Compute the scalar curvature of* $\tilde{g}$.

5 Electrovac solitons

The properties of electrovac solitons, that is, global stationary solutions of the Einstein–Maxwell vacuum equations, are proved along the same lines as the properties of gravitational solitons if one makes the additional hypothesis, often neglected in the literature, of the *global existence of an electromagnetic potential*, which is ensured only if M is diffeomorphic to R^n.

Remark 5.1 Topology of M is studied in Chruściel and Wald[25].

Theorem 5.2 *Under the hypotheses made in Theorem 4.2 or Corollaries 4.3, and 4.4 on a stationary spacetime the conclusions of this theorem or corollaries hold for electrovac solitons with global electromagnetic potential A, under the hypothesis in the asymptotically Euclidean case that A_0 is bounded while A_i and F_{ij} fall off at infinity respectively like r^{-p} and $r^{-(p+1)}$, with $p > \frac{n-2}{2}$.*

Proof If the electromagnetic field is stationary and admits an electromagnetic potential A then

$$F_{i0} = \partial_i A_0. \tag{5.1}$$

Using the value of the connection coefficients $\hat{\omega}^h_{ij} = \omega^h_{ij}$ and $\omega^0_{0j} = \psi^{-1}\partial_j\psi$ one sees that the second set of Maxwell equations gives in a stationary spacetime

$$\hat{\nabla}_\alpha F^{\alpha 0} \equiv -\psi^{-1}\nabla^i(\psi^{-1}\partial_i A_0) \equiv -\psi^{-Z}(\Delta_g A_0 - \psi^{-1}\partial_i\psi\partial_i A_0) = 0 \tag{5.2}$$

$$\hat{\nabla}_i F^{ij} \equiv \nabla_i F^{ij} = 0. \tag{5.3}$$

Equation (5.2) implies, under our hypotheses, $A_0 =$ constant by the maximum principle, hence $F_{i0} \equiv 0$.

Having assumed the global existence of the space vector potential A_j we deduce from Equation (5.3) the global equation on M

$$A_j\nabla_i F^{ij} \equiv \nabla_i(A_j F^{ij}) - (\nabla_i A_j)F^{ij} = 0; \tag{5.4}$$

[25] Chruściel, P. and Wald, R. (1994) *Class. Quant. Grav.*, **11** L147–152.

that is, by the definition of the vector potential,

$$\nabla_i(A_j F^{ij}) = \frac{1}{2} F_{ij} F^{ij}. \tag{5.5}$$

The conclusion $F_{ij} \equiv 0$ follows by the same argument as in the proof of Theorem 3.2.

The spacetime is therefore vacuum Einstein. We apply Theorem (4.2). □

Remark 5.3 *When the "magnetic" 2-form $F_{ij} = 0$, then the momentum ρ_{0i} of the source vanishes, the spacetime is locally static by Theorem 3.2.*

6 The Einstein–Yang–Mills case

The previous proof does not extend to the Yang–Mills sources, as was remarked long ago by S. Deser[26]. The non-existence of **Einstein–Yang–Mills solitons** when $n = 3$ remained an open question until Bartnik and MacKinnon discovered numerically a static, spherically symmetric, complete and non-trivial solution of the Einstein–Yang–Mills equations.

7 Stationary black holes

It is believed that very massive objects will generally undergo gravitational collapse and give rise to a black hole, but this black hole will settle down to a stationary state.

It was conjectured already in the early 1940s by G. Darmois, without precise hypotheses, that the Schwarzschild spacetime was the only relevant one for vacuum static Einsteinian spacetimes. The uniqueness of general stationary black holes was predicted in the early 1950s by J. A. Wheeler using the picturesque phrase "Black holes have no hair". We will discuss the status of these conjectures in the last subsections. We will try to limit the exposition to fairly simple concepts, or definitions previously used in this book. More subtle and technical recent considerations relative to black holes can be found elsewhere.

We have seen that there are no $3 + 1$ gravitational solitons with compact in space support and sources with non-negative energy. There is no theorem on the non-existence of spacetimes with compact space outside of the black hole region, but no examples either of such spacetimes. In all this chapter the complements of the black hole regions are asymptotically Euclidean manifolds with boundary.

7.1 *Definitions*

We have recalled in Appendix VI the definition often used in the modern literature of black holes by conformal compactification for asymptotically Euclidean black holes, and we gave an analogous definition, without using the conformal compactification, in Chapter 13.

[26] Who encouraged the author of this book to look for a proof, which she did not find (fortunately, since it would have been false).

Black holes are strange objects in which space and time eventually become interchanged. A stationary black hole spacetime is initially defined as follows. More precision is given in Sections 8 and 11.

Definition 7.1 *A black hole spacetime is called stationary if it admits a Killing field which is timelike outside of the black hole region and preserves the horizon.*

Remark that no assumption is made on the character of the Killing field on the horizon (see Section 8).

7.2 *Axisymmetry*

Axisymmetric spacetimes are a particular case of spacetimes which admit a one-parameter spacelike isometry group. However, the group is not transitive: it admits a one-dimensional manifold of fixed points, the axis of symmetry. They are used to represent rotating bodies.

Physical considerations lead to an important conjecture formulated by Hawking and proved by him under an analyticity hypothesis and geometric assumptions: the conjecture that a stationary black hole is necessarily axisymmetric[27].

7.2.1 *Definitions*

We first give an elementary and intuitive definition.

Definition 7.2 *A sliced spacetime* $(V = M \times R, \hat{g})$ *is called axisymmetric if there exists in* M *a line (i.e. a one-dimensional submanifold)* A*, called the axis, and a group*[28] S^1 *of isometries of* $(V, \hat{g})$ *leaving invariant the manifolds* $M \times \{t\}$ *and leaving fixed* $A \times \{t\}$ *pointwise.*

One can imagine a spacetime both stationary and axisymmetric, but where the axis of symmetry is not invariant under the timelike isometries (i.e. the slicing considered above is not adapted to the timelike isometries). The following definition excludes such situations.

Definition 7.3 *A spacetime which is both stationary and axisymmetric is called a* ***stationary axisymmetric*** *spacetime if and only if its two Killing vector fields commute.*

We give a more mathematically elaborate definition of a stationary axisymmetric spacetime.

Definition 7.4 *A stationary axisymmetric* d*-dimensional spacetime* $(V \equiv M \times R, \hat{g})$ *is invariant under an Abelian group* $G_2 \equiv R \times S^1$*. It is the union of the product* $A \times R$ *of a spacelike line (the axis* A*) by* R *(timelike) with a principal fibre bundle* $V - \{A \times R\}$ *with Abelian group* $G_2 \equiv R \times S^1$ *acting on* $(V, \hat{g})$ *by timelike (under* R*) and spacelike (under* S^1*) isometries. The base* S*, the quotient*

[27] A non-axisymmetric rotating body would emit gravitational radiation, as pointed out by T. Damour, and hence cannot be stationary.

[28] Recall that S ^{1} ≡U(1)≡ SO(2).

of $V - \{A \times R\}$ by G_2, is a $d-2$-dimensional Lorentzian manifold. The axis A is pointwise invariant under the S^1 isometries.

Stationary **circular spacetimes** are particular cases of stationary axisymmetric ones which we define below.

Definition 7.5 *A stationary axisymmetric spacetime is circular if the 2-planes orthogonal to the two Killing fields are integrable; that is, are tangent to 2-surfaces.*

By this definition, in a circular d dimensional spacetime $(V, \hat{g})$ the manifold V is locally a product $\Sigma \times U$ where Σ is a 2-surface and U an open subset orthogonal to Σ of the $d-2$-dimensional manifold S.

7.2.2 *Stationary axisymmetric spacetime metrics*

The general stationary axisymmetric spacetime metric reads, in a local trivialization of V, according to the general Kaluza–Klein formula

$$\hat{g} = g_S + \Phi_{mn}(dx^m + A^m_a dx^a)(dx^m + A^m_a dx^a), \tag{7.1}$$

where g_S is a Riemannian metric on the basis S, reading in local coordinates

$$g_S \equiv g_{ab} dx^a dx^b, \quad a, b = 2, \ldots, d-2, \tag{7.2}$$

and Φ is a quadratic form of Lorentzian signature,

$$\Phi \equiv \Phi_{mn}(dx^m + A^m_a dx^a)(dx^m + A^m_a dx^a), \quad m, n = 0, 1 \tag{7.3}$$

$x^0 \equiv t$, coordinate on R, $x^1 = \phi$, coordinate on S^1. The coefficients Φ_{mn}, A^m_a depend only on the x^a.

7.2.3 *Stationary circular spacetime metrics*

The general stationary circular spacetime metric reads, in a subset $U \times \Sigma$ under the form (7.1), but with $A^m_a \equiv 0$. The quadratic form Φ can then be written as a metric on $S^1 \times R$,

$$\Phi \equiv -N^2 dt^2 + 2a dt d\phi + b^2 d\phi^2, \tag{7.4}$$

with coefficients N, a, b depending only on the coordinates x^a of the basis S.

Remark 7.6 *The metric of a stationary circular spacetime is invariant under a simultaneous time and angle of rotation reversals, $t \to -t$, $\phi \to -\phi$.*

A stationary circular spacetime is called static if the timelike orbits and the spacelike orbits of the respective one-parameter isometry groups R and S^1 are orthogonal, the coefficient a in (7.4) being zero.

8 The rigidity theorem for black holes

Section contributed by James Isenberg

Well-known explicit black hole solutions (Schwarzschild, Kerr, Kerr–Newman) are both stationary and axisymmetric. Presuming that we restrict our attention to equilibrium and therefore stationary black holes, one may ask if it is necessary that such solutions be axisymmetric as well. Physical arguments – that a non-axisymmetric solution would radiate away any multipole moments of order higher than that of angular momentum – suggest that axisymmetry does necessarily hold. In this section, we briefly describe the mathematical work which proves this result[29].

The starting point is the definition of a stationary black hole. The usual definition is that an $n+1$-dimensional spacetime (V, g) contains a black hole if it is asymptotically flat in the sense of admitting a conformal asymptotic structure including a future null infinity $\mathcal{I}^+$, and if the complement with respect to V of the causal past $J^-(\mathcal{I}^+)$ of $\mathcal{I}^+$ is non-empty. This complement region $V \setminus J^-(\mathcal{I}^+)$ of V is called the black hole region, and the boundary of the black hole, $\mathcal{H} := \partial(V \setminus J^-(\mathcal{I}^+))$, is the event horizon. This horizon, according to Hawking and Ellis and according to Wald is a Lipschitz embedded co-dimension one null hypersurface. Under some further hypotheses the horizon $\mathcal{H}$ admits a foliation by null geodesic paths (called "generators" of the horizon).

Some researchers are unhappy with this definition, and others have questioned details of the conditions under which $\mathcal{H}$ has a smooth foliation by null generators. However, for the purposes of proving rigidity, this structure is key, so we shall consider the case in which the spacetime containing a black hole does have a smooth horizon with null geodesic generators.

The notion of stationarity has been discussed above: a stationary black hole spacetime admits a Killing field T which is asymptotically (in a neighbourhood of $\mathcal{I}^+$) timelike, and which has complete orbits. We note that since the action on a spacetime generated by a Killing field must preserve isolated geometric structures, T must be tangent to the horizon $\mathcal{H}$. If, in addition, T is assumed to be tangent to the null geodesic generators of the horizon, then it would follow that T would necessarily be null and therefore normal to $\mathcal{H}$. The spacetime would then be static[30]. Since the results we discuss here are most useful if T is not tangent to the generators, we generally assume that the spacetime is stationary and not static.

At present it is not known if it is enough that a black hole spacetime be stationary for one to be able to show that it necessarily contains another Killing field. The results we discuss here require three additional conditions: first, it is necessary that the spacetime manifold, the metric, and the embedding of the

[29] See Hawking, S. and Ellis, G. (1973) *The Large Scale Structure of Space–Time*, Cambridge University Press and Chruściel, P. (2002) in Springer Lecture Notes in Physics 604, pp. 61–102, gr-qc 0201053 for further details.

[30] Sudarsky, D. and Wald, R. (1993) *Phys. Rev. D*, **12**(3), 5209–13.

horizon all be real analytic[31]. Second, we require that the generators of the horizon be geodesically incomplete in one direction. We need only make this assumption (which is called "non-degeneracy") on a single generator, since one can prove that if one generator is incomplete, then all of them must be. The third restriction we need to impose is topological. It is required that the event horizon (or the connected component of $\mathcal{H}$ being considered) be diffeomorphic to $\Sigma \times R$ for some compact $(n-1)$-dimensional manifold Σ, and in addition it is required that there be an embedding $i(\Sigma)$ of Σ in $\mathcal{H}$ such that the stationarity Killing field T is transverse[32] to $i(\Sigma)$.

With these assumptions clarified, we are now ready to state the main rigidity theorem:

Theorem 8.1 *Let (V, g) be an analytic, asymptotically flat black hole spacetime which is stationary (not static) and which satisfies the vacuum Einstein equations. Assume that there is a component $\mathcal{H}$ of the black hole horizon which is diffeomorphic to $\Sigma \times R$ for compact Σ and which has an analytic embedding of Σ in $\mathcal{H}$ which is transverse to the stationarity Killing field T. Then the spacetime admits a Killing field which is tangent to the null geodesic generators of $\mathcal{H}$ and everywhere independent of T.*

The proof of this result for $3+1$-dimensional black holes goes back to Hawking and Ellis, with a further refinement by Chruściel. In 2006 and 2007, using two different methods, Hollands, Ishibashi and Wald[33], and independently Moncrief and Isenberg[34] (MI2) have proven this theorem for general dimensions. The basic ideas behind the Chruściel and the Moncrief–Isenberg proofs are similar. We sketch them here.

The key first step is to use the action of the Killing field T to replace the given black hole spacetime by a new spacetime $(\tilde{V}, \tilde{g})$ with a compact null hypersurface $\tilde{\mathcal{H}}$. This is done by choosing a cross-section $i(\Sigma)$ of the black hole horizon $\mathcal{H}$ which is transverse to T, extending this cross-section into a spacetime codimension one hypersurface $i(\Sigma) \times R$ which is transverse to T as well, then acting on the spacetime via a diffeomorphism $\mathcal{A}$ generated by the Killing field, and finally identifying the spacetime at $i(\Sigma) \times R$ with $\mathcal{A}(i(\Sigma) \times R)$. This identification produces an analytic spacetime $(\tilde{V}, \tilde{g})$ which satisfies Einstein's equations and contains a compact null hypersurface $\tilde{\mathcal{H}}$ with a foliation by null geodesic generators.

[31] Chruściel, P., Delay, E., Galloway, G., and Howard, J. (2001) *Ann. Inst. Henri Poincaré*, **2**(1), 109–78 show that this third property is a consequence of the two previous ones.

[32] Note that if the spacetime contains a past null infinity $\mathcal{I}^-$ and if the horizon is contained in the future of $\mathcal{I}^-$, then (recalling that the vector field T is timelike near $\mathcal{I}^-$ and $\mathcal{I}^+$), it follows from an argument sketched in Chruściel and Wald (1994; *loc. cit.*) (see the proof of Proposition 4.1 in that reference), that one doesn't need to impose the requirement that T be transverse to $i(\Sigma)$ everywhere; it is automatic.

[33] Hollands, S., Ishibashi, A., and Wald, R. (2007) *Commun. Math. Phys.*, **271**(3), 699–722.

[34] Moncrief, V. and Isenberg, J. (1983) (2008) *Class. Quant. Grav.* **25**, 19015 (37 pp).

For the $3+1$-dimensional case, Σ must be a 2-sphere, and it follows from the topology of S^2 that one can always choose the identifying diffeomorphism $\mathcal{A}$ in such a way that the generators of the compactified horizon $\tilde{\mathcal{H}}$ are closed. Hence, as pointed out by Chruściel, one can apply the earlier Moncrief–Isenberg results[35] (MI1) to $(\tilde{V}, \tilde{g})$ to show that the spacetime must contain a Killing field K tangent to the generators of $\tilde{\mathcal{H}}$. Since the original spacetime is presumed to not be static, the (identified) stationarity Killing field $\tilde{T}$ is not tangent to these generators, and one concludes K and $\tilde{T}$ are linearly independent. The new Killing field of course survives the unwrapping of the spacetime $(\tilde{V}, \tilde{g})$ back to (a subset of) the black hole spacetime. One may argue further that linear combinations of the two vector fields produce a rotational Killing field on the black hole spacetime.

For a black hole spacetime of dimension $n+1$ with $n \geq 4$ the same operation can be carried out to produce a solution of Einstein's equations with a compact null hypersurface. The difference, however, is that one cannot generally do this in such a way that the generators on the compact null surface close. Hence one cannot simply apply the results of MI1 to obtain a Killing field tangent to the generators. In MI2[34] a generalization of the results from MI1 is carried out: The Killing field tangent to the generators of the null surface is produced, even though the generators are non-closed (so long as the Killing field T is present). The rest of the argument then proceeds as in the $3+1$ case.

One would like to remove the analyticity assumption, which to date is needed in the hypothesis of these black hole rigidity theorems. It is not yet clear if this can be done.

9 The Kerr metric and black hole

The stationary circular $3+1$-dimensional metric found by Kerr in 1963 plays a great role in General Relativity. Though no complete solution on $R^3 \times R$ with physically reasonable sources has been found to coincide with the Kerr solution in the exterior of a tube $\Omega \times R$ with Ω a compact set in R^3, it is generally believed that the asymptotic final state of a rotating body is represented (in the absence of fields other than gravitation) by a Kerr spacetime.

9.1 *Kerr metric in Boyer–Linquist coordinates*

The Kerr metric reads in coordinates found by Boyer and Linquist

$$g_{BL} = A\left(\frac{dr^2}{B} + d\theta^2\right) - \left(1 - \frac{2mr}{A}\right)dt^2 - \frac{4mar\sin^2\theta}{A}d\phi dt + \frac{2ma^2\sin^4\theta}{A}d\phi^2, \tag{9.1}$$

[35] Moncrief, V. and Isenberg, J. (1983) *Commun. Math. Phys.*, **89**(3) 489–513.

with m and a constants, t a time coordinate on R, and r, θ, ϕ interpreted as polar coordinates on R^3. The functions A and B are given by

$$A := r^2 + a^2 \cos^2 \theta, \quad B := r^2 - 2mr + a^2. \tag{9.2}$$

This stationary circular metric is smooth (analytic) and Lorentzian on the manifold

$$V_{BL} := \{R^3 - \Omega\} \times R, \quad \text{with} \quad R^3 - \Omega := \{A > 2mr, \quad B > 0\}. \tag{9.3}$$

Straightforward computation shows that (V_{BL}, g_{BL}) is a vacuum Einsteinian spacetime.

9.2 *The Kerr–Schild spacetime*

The Kerr metric g_{BL} in Boyer–Linquist coordinates is singular on various submanifolds of $R^3 \times R$. Like the Schwarzschild metric it is possible to extend it to a larger domain of $R^3 \times R$ than V_{BL}.

The following coordinates have been introduced by Kerr and Schild (1965). They are no more adapted to stationary circularity, but they permit the extension of the BL metric. The coordinate v, adapted to light propagation is such that

$$dv = dt + \frac{a^2 + r^2}{B}. \tag{9.4}$$

The new angular coordinate Φ is such that

$$d\Phi = d\phi + \frac{a}{B} dr. \tag{9.5}$$

The Boyer–Linquist metric takes then the Kerr–Schild form

$$\begin{aligned} g_{KS} \equiv &- \left(1 - \frac{2mr}{A}\right) dv^2 + 2drdv + Ad\theta^2 + \frac{\Sigma^2}{A} \sin^2 \theta d\Phi^2 \\ &+ 4amr \frac{\sin^2 \theta}{A} dv d\Phi + 2a \sin^2 \theta dr d\Phi \end{aligned} \tag{9.6}$$

with

$$\Sigma^2 \equiv (r^2 + a^2)^2 - Ba^2 \sin^2 \theta \tag{9.7}$$

This metric extends to a regular Lorentzian metric for all $A > 0$; it reduces to the Minkowski metric in retarded time and polar coordinates of R^3 if $m = a = 0$, and it reduces to the Eddington–Finkelstein metric when $a = 0$.

9.2.1 *Essential singularity*

The Kerr–Schild metric becomes singular for $A \equiv r^2 + a^2 \cos^2 \theta = 0$, i.e. $r = 0$, $\theta = \frac{\pi}{2}$. This is a genuine singularity, in that it can be checked that the **Kretschman**[36] scalar tends to infinity when A tends to zero.

[36] Square of the curvature tensor.

In order to have the Kerr–Schild metric reducing to a form of the Minkowski metric when $m = 0$ and $a \neq 0$ the coordinates r and θ are reinterpreted as follows: θ and Φ are coordinates on S^2, but $r = 0$ is not one point of R^3; it is assumed instead that the manifold R^3 defined by the cartesian coordinates (x, y, z) is represented in the variables (r, θ, Φ) through the mapping (oblate polar "coordinates"):

$$x = (r^2 + a^2)^{\frac{1}{2}} \sin\Phi \sin\theta, \quad y = (r^2 + a^2)^{\frac{1}{2}} \cos\Phi \sin\theta, \quad z = r\cos\theta$$

The singularity $A = 0$, i.e. $r = 0, \theta = \frac{\pi}{2}$, is then represented by the circle.

$$z = 0, \quad x^2 + y^2 = a^2$$

9.2.2 *Horizons*

1. Case $|a| > m$, this would concern very fast rotating bodies. We than have $B := r^2 - 2mr + a^2 > 0$ for all r, no surface $r =$constant is a null surface, the essential singularity is naked. This case is considered as physically unrealistic.

2. Case $|a| < m$. The Boyer-Linquist metric appears then as singular for $r = r_+$ or $r = r_-$, solutions of $B = 0$ given by

$$r_\pm = m \pm \sqrt{m^2 - a^2}.$$

When a tends to zero the BL metric tends to the Schwarzschild metric, r_+ tends to $2m$ and r_- tends to zero. The surfaces $r = r_\pm$ are not singular in the Kerr–Schild spacetime, but they are null surfaces, since the contravariant component g^{rr} in the Kerr–Schild metric vanishes for $B = 0$, as can be foreseen and checked by direct calculus.

The surface $r = r_+$ is the event horizon: no particle entering the region $r < r_+$ can escape from it. The future light cone at points where $r = r_+$ points entirely towards the interior (the vectors $k^0 = 1, k^1 = 0$ and $k^0 = 1, k^1 = -1$ are the two radial future pointing null vectors). The surface $r = r_-$ has no obvious physical meaning.

9.2.3 *Limit of stationarity. The ergosphere*

We see in Equation (9.6) that the variable r becomes a time variable where

$$\frac{2mr}{A} > 1$$

Then the metric is no longer stationary. The surface $r = r_{stat}$, with r_{stat} the largest root of $A - 2mr = 0$, is called the *limit of stationarity*. It does not coincide with the horizon if $a \neq 0$: it holds that

$$r_{stat} = m + \sqrt{m^2 - a^2\cos^2\theta} \geq r_+.$$

The domain between the limit of stationarity and the horizon r_+ is traditionally called the **ergosphere**[37].

[37] Though the name "ergoregion" would be more appropriate.

Proposition 9.1 ***(Penrose process)*** *A particle entering the ergosphere can leave it with a greater energy.*

Proof Consider the Killing vector $X \equiv \frac{\partial}{\partial v}$, with components (X^α) =(1,0,0,0) in Kerr–Schild coordinates. This vector is timelike outside of the ergosphere $r > r_{Stat}$,

$$g_{\alpha\beta}X^\alpha X^\beta \equiv -\left(1 - \frac{2mr}{A}\right) < 0 \quad \text{when} \quad r > r_{Stat},$$

and spacelike when $r_+ < r < r_{stat}$. A particle with timelike 4-momentum p has an energy $E = -p^\alpha X_\alpha$ with respect to the vector field X, this scalar is constant along the trajectory (a geodesic) because of the law of dynamics and the fact that X is a Killing vector. It holds that $E > 0$ outside the ergosphere. Suppose the particle enters the ergosphere and splits there into two particles with momenta p_1 and p_2. By the conservation law of the 4-momentum it holds that:

$$p = p_1 + p_2, \quad E = E_1 + E_2$$

with $E_i = -p_i^\alpha X_\alpha$. Since X is space like inside the ergosphere it is possible to have the timelike vector p_1 such that $E_1 < 0$, hence $E_2 = E - E_1 > E$. Both scalars E_1 and E_2 are conserved along the trajectories of the respective fragments. If the second one returns to the outside of the ergosphere it will be with a greater energy (with respect to the X observer) than the whole piece which left it. It is considered that it has extracted energy from the rotating black hole. □

9.2.4 *Extended Kerr spacetime*

The Kerr–Schild spacetime that we have defined is not complete. It is possible to extend it (in particular to negative r), but the results are much more complicated than in the Schwarzschild case and their physical interpretation very unclear, though geometrically interesting (cf. for instance Wald (1984), Chandrasekar (1983)[38].)

10 Uniqueness of stationary black holes (dimension 3 + 1)

We give in this section only a brief summary of basic results. They have been improved over the years by more and more elaborate proofs[39].

The next section contributed by P. Chruściel gives examples of other black holes, when the hypotheses of the uniqueness theorem are not satisfied, connectedness of the horizon in the first subsection, dimension 3 + 1 in the second subsection. There are many other interesting results for black holes with non-usual sources (Yang–Mills etc.), or for higher dimensional black holes.

[38] Wald, R. (1984) *General Relativity*, University of Chicago Press; Chandrasekhar, S. (1983) *The Mathematical Theory of Black Holes*, Clarendon Press, Oxford.

[39] See the review article by Chruściel (2002) *loc. cit.* and references therein.

10.1 *Static black holes*

The **Israel theorem**, uniqueness of the Schwarzschild black hole as a static asymptotically Euclidean black hole in vacuum was proved by Israel (1967) under the hypothesis that the 2-surfaces $\xi^i\xi_i =$ constant are smooth and homeomorphic to nested 2-spheres and the apparent horizon also is topologically spherical. He extended his proof to the Reissner–Nordström metric in 1968. Successive improvements using weaker hypotheses were obtained by various authors, in particular by D. Robinson who proved directly the vanishing of the Coton tensor R_{ijk} of a static black hole, and by Buntig and Masood ul Alam who proved the conformal flatness of the space by using the positive mass theorem. The hypothesis on the spherical nature of the horizon has been partially removed by Hawking, supposing it is simply connected and using the Euler Poincaré characteristic.

10.2 *Axisymmetric black holes*

The theorem for static black holes is a particular case (with $a = 0$) of the uniqueness theorem of the Kerr–Schild spacetime among asymptotically Euclidean stationary axisymmetric black holes in vacuum. The following **Carter–Robinson theorem** has been proved by Robinson (1975), improving previous results of Carter (1972).

Theorem 10.1 *Asymptotically Euclidean, stationary, axisymmetric spacetimes solutions of the vacuum Einstein equations which have a regular and connected event horizon, with a non-zero surface gravity*[40]*, are uniquely specified by two parameters, the mass m and the angular momentum a, if these spacetimes are smooth outside of the horizon and they are asymptotically flat.*

Since the Kerr metric provides a solution of the vacuum Einstein equations satisfying the hypothesis of the theorem for any given m and a, the Kerr metrics are the only such axisymmetric stationary black holes.

The simple connectedness and the spherical topology of the horizon has now been proved by Chruściel and Wald[41], under a global hyperbolicity hypothesis.

The uniqueness of the Kerr–Newman black hole in the electrovac case (with the restriction $m^2 > a^2 + Q^2$) has been proved by Mazur (1982).

10.3 *Uniqueness of the Kerr black hole*

We have seen in Section 8 that AF stationary black holes have been proven to be axisymmetric under analyticity assumptions. The removal of analyticity in the proof of the rigidity theorem is still an open question, interesting also in other contexts. The recent proof that the domain of outer communication of a smooth

[40] The surface gravity of a stationary black hole is defined in the next section. A physical interpretation can be found in Wald, section 12.5.

[41] Chruściel, P. and Wald, R. (1994) *Class. Quant. Grav.*, **11** L147–52.

four-dimensional stationary black hole is locally isometric to a Kerr spacetime by Ionescu and Klainerman[42] bypasses the axisymmetry property. They use for their proof a 4-tensor constructed through the Weyl tensor, called the Mars[43] tensor, the non-linear wave equations it satisfies and a new uniqueness theorem for this wave equation. The proof requires many geometrical and analytical ingredients and technical assumptions given in detail in the quoted paper.

11 Further results

Section contributed by P. Chruściel

11.1 *Multi black hole solutions*

The **Majumdar–Papapetrou** solutions provide a family of static metrics satisfying the Einstein–Maxwell equations with several black holes. The metric g and the electromagnetic potential A take the form[44] Compare Hartle, B. and Hawking, S. W. (1972) *Commun. Math. Phys.*, **26**, 87–101.

$$g = -u^{-2}dt^2 + u^2(dx^2 + dy^2 + dz^2)\,, \quad A = u^{-1}dt\,, \tag{11.1}$$

$$u = 1 + \sum_{i=1}^{I} \frac{\mu_i}{|\vec{x} - \vec{a}_i|}\,, \tag{11.2}$$

for some positive constants μ_i (the electric charges carried by the punctures $\vec{x} = \vec{a}_i$). *Standard MP black holes* are obtained if the coordinates of x^μ cover the range $\mathbf{R} \times (\mathbf{R}^3 \setminus \{\vec{a}_i\})$ for a finite set of points $\vec{a}_i \in \mathbf{R}^3$, $i = 1, \ldots, I$. It has been shown by Hartle and Hawking that standard MP spacetimes can be analytically extended to an electrovac spacetime with I black hole regions.

The case $I = 1$ is a special case of the so-called *Reissner–Nordström* metrics, which are the charged, spherically symmetric (connected) generalizations of the Schwarzschild black holes.

The static electrovac black holes are well understood, but in order to present the results some terminology is needed. A hypersurface (Σ, g, K) (perhaps with boundary) is called *asymptotically flat* if Σ is the union of a compact set and a region Σ_{ext} diffeomorphic to $\mathbf{R}^n \setminus B(R)$, for some ball of radius R, called an *asymptotic end*. On Σ_{ext} the metric g is assumed to approach the Euclidean metric at least as fast as $1/r$, with the first derivatives falling-off one power faster, while the extrinsic curvature tensor K approaches zero at least as fast as $1/r^2$. In electro-vacuum, the Maxwell field is furthermore assumed to fall off at least as fast as r^{-2}, with the first derivative again falling off one power faster.

A null hypersurface which coincides with a connected component of the set

$$N(X) := \{g(X, X) = 0\ ,\ X \neq 0\}\ ,$$

[42] Ionescu, D. and Klainermanz, S. (2007) gr.qc 0711.0040v1

[43] Mars, M. (2000) *Class. Quant. Grav.*, **17**, 335–73.

[44] Majumdar, S. D. (1947) *Phys. Rev.*, **72**, 390–8; Papapetrou, A. (1947) *Proc. R. Irish Acad.*, **A51**, 191–204.

where X is a Killing vector tangent to $N(X)$, is called a **Killing horizon** associated with X. The **surface gravity** κ of a Killing horizon is defined by the formula

$$d\,(g(X,X)) = -2\kappa X^\flat \ , \tag{11.3}$$

where $X^\flat = g_{\mu\nu}\, X^\nu dx^\mu$. A fundamental property is that the surface gravity κ is constant over each horizon in vacuum, or in electro-vacuum[45]. A Killing horizon is called a **degenerate horizon** if κ vanishes, and *non-degenerate* otherwise.

A key notion in the theory of black holes is that of the **domain of outer communications**: as previously, a spacetime (M, g) will be called quasi-stationary if there exists on M a Killing vector field X which is uniformly timelike in the asymptotic region $\Sigma_{\rm ext}$; X will be referred to as the *quasi-stationary* Killing vector. For $t \in \mathbf{R}$ let $\phi_t : M \to M$ denote the one-parameter group of diffeomorphisms generated by X. The *domain of outer communications* is then defined as

$$\langle M_{\rm ext}\rangle = J^+(\cup_t \phi_t(\Sigma_{\rm ext})) \cap J^-(\cup_t \phi_t(\Sigma_{\rm ext}))\ .$$

It can be shown that $\langle M_{ext}\rangle$ is independent of the choice of Σ_{ext}, provided that the resulting $\langle M_{ext}\rangle's$ overlap. The boundary of $\langle M_{\rm ext}\rangle$ is called the *event horizon*.

A quasi-stationary spacetime (M, g) is called I^+ *regular* if the domain of outer communications $\langle M_{ext}\rangle$ is globally hyperbolic. Moreover, it is assumed that M contains a connected, spacelike, acausal hypersurface Σ, the M-closure of which is a topological manifold with boundary, consisting of the union consisting of the union of a compact set and a finite number of asymptotically flat ends. The (compact boundary $\partial\bar{\Sigma} := \bar{\Sigma}\backslash\Sigma$ is assumed to be a topological manifold satisfying

$$\partial\bar{\Sigma} \subset E^\circ := \partial(\langle M_{\rm ext}\rangle) \cap I^+(\Sigma_{ext})$$

and intersecting every generator of E^+ exactly once.

A theorem of Carter and Wishweshwara asserts that *in regular black hole spacetimes event horizons are Killing horizons.*

The analysis in Carter (1973), Ruback (1988), and Somon (1992)[46] as completed in Chruściel (1999)[47] leads to:

Theorem 11.1 *Every domain of outer communication in a static, electro-vacuum, regular black hole spacetime without degenerate horizons is isometrically diffeomorphic to a domain of outer communications of a Reissner–Nordström black hole.*

[45] Heusler, M. (1996) *Black Hole Uniqueness Theorems*, Cambridge University Press.

[46] Carter, B. (1973) In *Black Holes* (eds. C. de Witt and B. de Witt), Gordon & Breach; Ruback, P. (1988) *Class. Quant. Grav.*, **5**, L155–9; Simon, W. (1992) *Class. Quant. Grav.*, **9**, 241–56.

[47] Chruściel, P. T. (1999) *Class. Quant. Grav.*, **16**, 689–704, arXiv:gr-qc/98100.

The relevance of the standard MP black holes follows now from the following result of Chruściel and Tod[48]:

Theorem 11.2 *Every domain of outer communication in a static, electro-vacuum, regular black hole spacetime containing degenerate horizons is isometrically diffeomorphic to a domain of outer communications of a standard MP spacetime.*

It thus follows that the MP family provides the only static, electro-vacuum, regular black holes with *non-connected* horizons.

11.2 *The Emparan–Reall "black rings"*

A fascinating class of black hole solutions of the 4+1-dimensional stationary vacuum Einstein equations has been found by Emparan and Reall[49]. The solutions are asymptotically Minkowskian in spacelike directions, with an ergosurface and an event horizon having $S^1 \times S^2$ cross-sections. (The "ring" terminology refers to the S^1 factor in $S^1 \times S^2$.) We include a detailed analysis, as an illustration of the methods at hand.

The spacetime is obtained by gluing together two coordinate patches, in each of which the metric takes the form

$$\begin{aligned} g = &-\frac{F(x)}{F(y)}\left(dt + \sqrt{\frac{\nu}{\xi_F}}\frac{\xi_1 - y}{A}d\psi\right)^2 \\ &+ \frac{F(y)}{A^2(x-y)^2}\left[-F(x)\left(\frac{dy^2}{G(y)} + \frac{G(y)}{F(y)}d\psi^2\right)\right. \\ &\left.+F(y)\left(\frac{dx^2}{G(x)} + \frac{G(x)}{F(x)}d\varphi^2\right)\right], \end{aligned} \tag{11.4}$$

where $A > 0$, ν, and ξ_F are constants, and

$$F(\xi) = 1 - \frac{\xi}{\xi_F}, \tag{11.5}$$

$$G(\xi) = \nu\xi^3 - \xi^2 + 1 = \nu(\xi - \xi_1)(\xi - \xi_2)(\xi - \xi_3). \tag{11.6}$$

The constant ν is chosen to satisfy $0 < \nu \leq \nu_* = 2/3\sqrt{3}$. The upper bound is determined by the requirement that the three roots $\xi_1 < \xi_2 < \xi_3$ of G are real, and in fact

$$\xi_1 < 0 < \xi_2 < \xi_3.$$

[48] Chruściel, P. T. and Tod, K. P. The classification of static electro-vacuum space–times containing an asymptotically flat spacelike hypersurface with compact interior, arXiv:gr-qc/0512043.

[49] Emparan, R. and Reall, H. S. (2002) *Phys. Rev. Lett.*, **88**, 101, arXiv:hep-th/0110260; (2004) *J. High Energy Phys.*, **3**, 64; (2006) *Class. Quant. Grav.*, **23**, R169–R197, arXiv:hep-th/0608012.

The choice

$$\xi_F = \frac{\xi_1\xi_2 - {\xi_3}^2}{\xi_1 - 2\xi_3 + \xi_2}$$

guarantees that the singularity of the metric at the sets $x = \xi_1$ and $x = \xi_2$ disappears, if one interprets those sets as axes of rotation, with φ being an angular-type coordinate around those axes. Then the coordinates

$$\tilde{\rho} = 2\sqrt{x - \xi_1}\ , \quad \varphi = \lambda\tilde{\varphi}\ , \tag{11.7}$$

where λ is given by

$$\lambda = \frac{2\sqrt{\xi_F - \xi_1}}{\nu\sqrt{\xi_F}(\xi_2 - \xi_1)(\xi_3 - \xi_1)}.$$

become polar coordinates centred on the north pole on S^2. One has

$$\xi_2 \leq \xi_F < \xi_3\ .$$

The requirement

$$\xi_1 \leq x \leq \xi_2 \tag{11.8}$$

leads to $G(x) \geq 0$ and $F(x) > 0$.

A Lorentzian signature is obtained away from zeros and infinities except if $F(y) > 0$ and $G(y) > 0$, leading to the following allowed values of y

$$y \in (-\infty, \xi_1] \cup (\xi_F, \infty). \tag{11.9}$$

Anticipating: the singular set $y = x = \xi_1$ corresponds to "the infinity of an asymptotically flat region". The "boundary $y = -\infty$" is a smooth ergosurface; similarly "$y = \infty$". Furthermore one can glue the regions $y < \xi_1$ and $y > \xi_F$ along this boundary obtaining a real analytic metric, with a Killing horizon at $y = \xi_3$.

Given the ranges of coordinates, the singularity $x = y$ occurs only for $x = y = \xi_1$. So, the coordinate t parametrizes $\mathbf{R}$, the coordinates (y, ψ) are related to polar coordinates $(\hat{\rho}, \hat{\varphi})$ on $\mathbf{R}^2$ by formulae analogous to (11.7), and for $y \neq \xi_1$ the coordinates (x, φ) are coordinates on S^2. If we think of $x = \xi_1$ as being the north pole of S^2, and we denote it by N, and if $y = \xi_1$ is thought of as the origin of $\mathbf{R}^2$, then g is an analytic metric on

$$\underbrace{\mathbf{R}}_{t} \times \left\{ \left(\underbrace{\mathbf{R}^2}_{y,\psi \Leftrightarrow \hat{\rho},\hat{\varphi}} \times \underbrace{S^2}_{x,\varphi \Leftrightarrow \tilde{\rho},\tilde{\varphi}} \right) \Big\backslash \left(\{0\} \times \{N\} \right) \right\}.$$

Near the singular set $\mathbf{R} \times \{0\} \times \{N\}$, Emparan and Reall replace $(\tilde{\rho}, \hat{\rho})$ by new polar variables $(\tilde{r}, \hat{r})$ defined as

$$\tilde{r} = \frac{\tilde{\rho}}{B(\tilde{\rho}^2 + \hat{\rho}^2)}\ , \quad \hat{r} = \frac{\hat{\rho}}{B(\tilde{\rho}^2 + \hat{\rho}^2)}\ , \tag{11.10}$$

Setting

$$\hat{y}^1 = \hat{r}\cos\hat{\varphi}\ , \quad \hat{y}^2 = \hat{r}\sin\hat{\varphi}\ , \quad \tilde{y}^1 = \tilde{r}\cos\tilde{\varphi}\ , \quad \tilde{y}^2 = \tilde{r}\sin\tilde{\varphi}\ ,$$

introducing $(x^\mu) = (t, \hat{y}^1, \hat{y}^2, \tilde{y}^1, \tilde{y}^2)$, and choosing B so that

$$(4B)^2 F^2(\xi_1) = \nu A^2(\xi_1 - \xi_2)(\xi_1 - \xi_3)\ ,$$

one obtains – rather surprisingly – a manifestly asymptotically flat metric:

$$g = \{\eta_{\mu\nu} + O(r^{-2})\}dx^\mu dx^\nu\ .$$

In order to understand the geometry when $y \to -\infty$, one replaces y by

$$Y = -1/y\ .$$

As yet another surprise, the metric can be analytically extended across $\{Y = 0\}$ to negative Y: indeed, we have

$$\begin{aligned} g = -F(x)\Bigg[&\frac{dt^2}{F(y)} + 2\sqrt{\frac{\nu}{\xi_F}}\frac{\xi_1 - y}{AF(y)}dtd\psi \\ &+\frac{1}{A^2}\left\{\frac{\nu(\xi_1 - y)^2}{\xi_F - y} + \frac{G(y)}{(x-y)^2}\right\}d\psi^2 + \frac{F(y)y^4}{A^2(x-y)^2G(y)}dY^2\Bigg] \\ &+ \frac{F^2(y)}{A^2(x-y)^2}\left(\frac{dx^2}{G(x)} + \frac{G(x)}{F(x)}d\varphi^2\right), \end{aligned} \tag{11.11}$$

hence when y tends to $-\infty$

$$\begin{aligned} &- F(x)\left[2\frac{\sqrt{\nu\xi_F}}{A}dtd\psi + \frac{2\xi_1 + \acute{e}x - 1 - \nu\xi_F}{A^2}d\psi^2 + \frac{1}{A^2\nu\xi_F}dY^2\right] \\ &+ \frac{1}{A^2{\xi_F}^2}\left(\frac{dx^2}{G(x)} + \frac{G(x)}{F(x)}d\varphi^2\right). \end{aligned} \tag{11.12}$$

One checks that the metric remains non-degenerate up to $\{Y = 0\}$. and that all functions in (11.11) extend analytically to small negative Y; e.g.

$$g(\partial_t, \partial_t) = g_{tt} = -\frac{F(x)}{F(y)} = -\frac{\xi_F - x}{\xi_F - y} = -\frac{(\xi_F - x)Y}{Y\xi_F + 1}. \tag{11.13}$$

To take advantage of the work done so far, in the region $Y < 0$ one can now replace Y by a new coordinate

$$z = -Y^{-1} > 0\ ,$$

obtaining a metric which has the same form as (11.4), with y there replaced by z. In view of the previous analysis, the metric is manifestly regular in the range

$$\xi_3 < z < \infty\ . \tag{11.14}$$

Note, however, that the "stationary" Killing vector ∂_t, which was timelike in the region $Y > 0$, is now spacelike. Therefore the region (11.14) is an *ergoregion* for the extended spacetime. The *ergosurface* at $Y = 0$ has topology $S^1 \times S^2$.

The above coordinates break down at $z = \xi_3$, where one replaces ψ by a new (periodic) coordinate χ, and t by a new coordinate v, defined as

$$d\chi = d\psi + \frac{\sqrt{-F(z)}}{G(z)} dz \,, \tag{11.15}$$

$$dv = dt + \sqrt{\frac{\nu}{\xi_F}} (z - \xi_1) \frac{\sqrt{-F(z)}}{AG(z)} dz \,. \tag{11.16}$$

In the new coordinates the metric takes the form

$$\begin{aligned} ds^2 = &- \frac{F(x)}{F(z)} \left(dv - \sqrt{\frac{\nu}{\xi_F}} \frac{z - \xi_1}{A} d\chi \right)^2 \\ &+ \frac{1}{A^2 (x - z)^2} \left[F(x) \left(-G(z) d\chi^2 + 2\sqrt{-F(z)} d\chi dz \right) \right. \\ &\left. + F(z)^2 \left(\frac{dx^2}{G(x)} + \frac{G(x)}{F(x)} d\phi^2 \right) \right] . \end{aligned} \tag{11.17}$$

This is regular at

$$E := \{ z = \xi_3 \} \,,$$

and the metric can be analytically continued into the region $\xi_F < z \leq \xi_3$. One can check directly from (11.17) that $g(\nabla z, \nabla z)$ vanishes at E, so that E is a null hypersurface. Furthermore, z is a time function on $\{ z < \xi_3 \}$, which is thus a black hole region by the usual arguments. The topology of the event horizon is $\mathbf{R} \times S^1 \times S^2$: this is a "rotating black ring".

XV

GLOBAL EXISTENCE THEOREMS ASYMPTOTICALLY EUCLIDEAN DATA

1 Introduction

The existence of past and future complete asymptotically Euclidean[1] vacuum Einsteinian spacetimes already preoccupied Einstein. He thought that these spacetimes being empty should coincide, in the absence of radiation coming from infinity, with the Minkowski spacetime. His physical intuition was correct, but it lacked a precise formulation of the "absence of radiation coming from infinity". The positive mass theorem first formulated by Deser, with a partly heuristic proof[2] by Brill and Deser, makes precise the strong fall-off behaviour at space infinity required of the spacetime metric for the conjecture of Einstein to hold.

The existence of past and future complete vacuum asymptotically Euclidean Einsteinian spacetimes, with the usual "Newtonian type" fall-off at space infinity has remained an open problem, with controversial conjectured solutions, until a monumental paper of D. Christodoulou and S. Klainerman, "Non-linear stability of Minkowski space", now published as a book [3], solved it in the affirmative. A little earlier H. Friedrich had proved, using a conformal formulation, global future evolution for small data on an hyperboloid which turns null at infinity[4].

In Section 2 we show how the Penrose transform can be used to prove global existence of solutions of various semilinear field equations; in Subsection 2.2 we give the proof for quasilinear wave equations in spacetime dimension greater or equal to 6, or in spacetime dimension 4 and satisfying the "null condition" found independently by Christodoulou and Klainerman.

In Section 3 we outline the foundation points of Friedrich's conformal system.

In Section 4 we explain how a conformal transformation of a future causal cone in Minkowski spacetime of dimension greater than or equal to six on to

[1] Recall that we call asymptotically Euclidean a spacetime (V, g) where $V = M \times R$, the sections $(M_t, \bar{g}_t)$ are asymptotically Euclidean, and on each M_t the lapse tends to 1 and the shift tends to zero at infinity.

[2] Made mathematically rigorous for a neighbourhood of Minkowski spacetime by Choquet-Bruhat and Marsden (see related paper "Positive mass theorem", published in DeWitt, C. and DeWitt, B. (eds.) (1983) *Relativity Groups and Topology*, Gordon and Breach).

[3] Christodoulou, D. and Klainerman, S. (1993), Princeton University Press.

[4] Friedrich, H. (1986) *Commun. Math. Phys.*, **107**, 587–609.

another such light cone gives a global existence theorem of solutions of the vacuum Einstein equations with small data which are Schwarzschild outside of a compact set.

In Sections 5 and 6 we indicate some of the arguments of the book *Non-Linear Stability of Minkowski Space*, and state some further properties proved in another book by Nicolo and Klainerman.

Finally, in Section 8 we sketch the main steps of the proof by Lindblad and Rodnianski of the global existence in wave coordinates, for small initial data.

2 Global existence for small data via the Penrose map

The Penrose confomorphism **P** (see Appendix VI) is the mapping **P** from R^{n+1} into $S^n \times R$ given by

$$T = tg^{-1}u + tg^{-1}v, \quad \alpha = tg^{-1}u - tg^{-1}v, \quad u =: t + r, \quad v =: t - r. \tag{2.1}$$

P is a diffeomorphism from the Minkowski spacetime $M^{n+1} := (R^{n+1}, \eta)$ onto the bounded open subset $\mathcal{V} \subset S^n \times (-\pi, \pi)$ of the Einstein cylinder $\Sigma^{n+1} := (S^n \times R, \gamma)$,

$$\mathcal{V} := \{0 \leq \alpha < \pi, \quad (\theta^i) \in S^{n-1}, \quad \alpha - \pi < T < \pi - \alpha\}, \tag{2.2}$$

with the conformal property

$$\mathbf{P}^* \eta \equiv \Omega^{-2}\gamma, \quad \Omega := (\cos T + \cos \alpha) \tag{2.3}$$

where η is the Minkowski metric, which reads in polar coordinates on R^n

$$\eta := -dt^2 + dr^2 + r^2(dS^{n-1})^2$$

and γ is the metric on the Einstein cylinder

$$\gamma \equiv -dT^2 + d\alpha^2 + \sin^2 \alpha (dS^{n-1})^2. \tag{2.4}$$

The submanifold R^n, $t = 0$, of M^{n+1} is mapped onto the sphere S^n, $T = 0$, of Σ^{n+1}, minus a point, $\alpha = \pi$, traditionally named I_0. If a differential system $\mathcal{S}$ is transformed by the Penrose map into a differential system which extends to a system $\tilde{\mathcal{S}}$ on Σ^{n+1} for which a Cauchy problem with data at $T = 0$ admits solutions up to time parameter $T = \pi$, then[5] the corresponding Cauchy problem with data at $t = 0$ for the original system $\mathcal{S}$ admits global solutions on R^{n+1}. Unfortunately the smoothness of the data on S^n required for the solution of the Cauchy problem for $\tilde{\mathcal{S}}$ imposes fall-off of the original data on R^n which may restrict the interest for physics. On the other hand, the existence of the solution of the transformed system up to time $T = \pi$ gives the asymptotic fall-off of the solution of the original system.

[5] Christodoulou, D. (1981) *C. R. Acad. Sci. Paris*, **293**, 79–42.

2.1 *Yang–Mills and associated equations*

The four-dimensional Yang–Mills equations are conformally invariant, the global existence for small data can be obtained by the Penrose confomorphism method. The result extends to the spinor and massless scalar multiplet equations[6].

The global existence of solutions of the Cauchy problem for the Yang–Mills equations on the Einstein cylinder has been proved[7] through the use of the L^∞ estimates obtained by Eardley and Moncrief on Minkowski spacetime M^4. The conformal mapping from M^4 onto half of the Einstein cylinder Σ^4 (see Appendix VI) has given a proof of global existence of solutions of Yang–Mills equations on anti-de Sitter spacetime[8].

2.2 *Quasi-linear wave equations*

The scalar curvatures of M^{n+1} and Σ^{n+1} are respectively

$$R(\eta) = 0 \quad \text{and} \quad R(\gamma) = n(n-1)$$

The identity, from Appendix VI, between wave operators of two conformal metrics reads in this case, with $\tilde{g} = \eta$, in arbitrary coordinates,

$$\Box_\gamma u - \frac{(n-1)^2}{4} u \equiv \Omega^{-(3+n)/2} \Box_\eta \tilde{u}, \quad \tilde{u} = \Omega^{(n-1)/2} u. \tag{2.5}$$

We consider the quasilinear second-order equation on R^{n+1} with unknown a scalar function $\tilde{u}$, written in a global system of coordinates (x^α) of R^{n+1}, with ∂_α the partial derivative and $\tilde{g} := \eta$ the Minkowski metric.

$$F(\tilde{u}, \partial\tilde{u}, \partial^2\tilde{u}) := \{\eta^{\alpha\beta} + h^{\alpha\beta}(\tilde{u}, \partial\tilde{u})\}\tilde{\nabla}_\beta \partial_\alpha \tilde{u} + f(\tilde{u}, \partial_\alpha \tilde{u}) = 0. \tag{2.6}$$

We suppose that the function f depends smoothly on its arguments.

The conformal law (2.5) shows that Equation (2.6) for $\tilde{u}$ on M^{n+1} is equivalent to the equation for $u = \Omega^{-(n-1)/2}\tilde{u}$ on the bounded set $\mathcal{V} \subset \Sigma^{n+1}$,

$$\begin{aligned} \gamma^{\alpha\beta}\nabla_\beta \partial_\alpha u - \frac{(n-1)^2}{4} u + \Omega^{-(3+n)/2}\{h^{\alpha\beta}(\tilde{u}, \partial\tilde{u})\}\tilde{\nabla}_\beta \partial_\alpha \tilde{u} \\ + f(\tilde{u}, \partial_\alpha \tilde{u})\} = 0, \quad \tilde{u} = \Omega^{-(n-1)/2} u. \end{aligned} \tag{2.7}$$

Christodoulou[9] has used this observation to obtain global solutions on Minkowski spacetime of the Cauchy problem with small initial data for classes of quasilinear wave equations where h and f, at least in a neighbourhood of $\tilde{u} = 0$, are smooth functions of $\tilde{u}$ and $\partial\tilde{u}$. Christodoulou assumes that, for $\tilde{u} = 0$, it holds that:

$$h(0,0) = 0, \quad f(0,0) = 0, \quad \frac{\partial f}{\partial \tilde{u}}(0,0) = 0, \quad \frac{\partial f}{\partial(\partial\tilde{u})}(0,0) = 0. \tag{2.8}$$

[6] Choquet-Bruhat, Y. and Christodoulou, D. (1981) *Ann. Sci. E*; N.S. serie 4, **14**, 481–500.
[7] Choquet-Bruhat, Y., Paneitz, S., and Segal, I. E. (1983) *J. Funct. Anal.*, **53**(2) 112–50.
[8] Choquet-Bruhat, Y. (1989) *Class. Quant. Grav.*, **6**, 1781–6.
[9] Christodoulou, D. (1986) *Commun. Pure Appl. Math.*, **39**, 267–82.

We will give a variant of the general proof given by Christodoulou, considering the case where the equations have the form taken by usual field equations in appropriate gauges, in particular by Einstein equations in wave coordinates; that is, h does not depend on $\partial\tilde{u}$ and f is a quadratic form in $\partial\tilde{u}$ with coefficients depending only on u.

We spell out in components the quasidiagonal hyperquasilinear system of wave equations that we consider

$$(\eta^{\alpha\beta} + h^{\alpha\beta}(\tilde{u})\}\tilde{\nabla}_\alpha\partial_\beta\tilde{u}^A + Q_{BC}^{\alpha\beta,A}(\tilde{u})\partial_\alpha\tilde{u}^B\partial_\beta\tilde{u}^C = 0. \tag{2.9}$$

Recall that $\tilde{g}$ is the Minkowski metric. The hypotheses (2.8) are then satisfied as soon as $h^{\alpha\beta}(0) = 0$.

Theorem 2.1 *When $h(0) = 0$ the Cauchy problem for the quasilinear quasidiagonal system of wave equations (2.9) with initial data on R^n has a global solution on M^{n+1}, with $n \geq 5$ and odd, if the image of the initial data by the Penrose map is in $H_s \times H_{s-1}(S^n)$, $s > \frac{n}{2} + 1$, and small enough in the corresponding norm.*

Proof We denote by x^α the Minkowskian coordinates t, r, θ^i; they are also local coordinates in the open set $\mathcal{V}$ of Σ^{n+1}, defined by the Penrose diffeomorphism. We denote by X^A the global coordinates T, α, θ^i on Σ^{n+1}. We still denote by $\tilde{\nabla}$ and ∇ covariant derivatives in respectively the Minkowski and the Einstein cylinder metrics. The system (2.9) on M^{n+1} is equivalent to the system on $\mathcal{V}$ which reads in Minkowskian coordinates

$$\gamma^{\alpha\beta}\nabla_\beta\partial_\alpha u - \frac{(n-1)^2}{4}u + \Omega^{-(3+n)/2}\{h^{\alpha\beta}(\Omega^{(n-1)/2}u)\tilde{\nabla}_\alpha\partial_\beta(\Omega^{(n-1)/2}u)$$
$$+ Q(\Omega^{(n-1)/2}u)[\partial(\Omega^{(n-1)/2}u), \partial(\Omega^{(n-1)/2}]\} = 0. \tag{2.10}$$

The operator $\gamma^{\alpha\beta}\nabla_\beta\partial_\alpha$ is invariant under diffeomorphisms; in arbitrary coordinates on $\mathcal{V}$, in particular the coordinates X^A, it holds that

$$\gamma^{\alpha\beta}\nabla_\beta\partial_\alpha u = \gamma^{AB}\nabla_B\partial_A u \tag{2.11}$$

This term extends to the whole of Σ^{n+1}. The non-linear term of (2.12) is more subtle to express as a function of $u, \partial_A u$ and $\nabla_A\partial_B u$. We deduce from the relations (2.1) that

$$\frac{\partial t}{\partial T} = \Omega^{-2}(1 + \cos T\cos\alpha), \quad \frac{\partial t}{\partial\alpha} = \Omega^{-2}\sin T\sin\alpha \tag{2.12}$$

$$\frac{\partial r}{\partial T} = \Omega^{-2}\sin T\sin\alpha, \quad \frac{\partial r}{\partial\alpha} = \Omega^{-2}(1 + \cos T\cos\alpha). \tag{2.13}$$

The components of the inverse matrix are

$$\frac{\partial T}{\partial t} = 1 + \cos T\cos\alpha, \quad \frac{\partial\alpha}{\partial t} = -\sin T\sin\alpha, \tag{2.14}$$

$$\frac{\partial T}{\partial r} = -\sin T\sin\alpha, \quad \frac{\partial\alpha}{\partial r} = 1 + \cos T\cos\alpha. \tag{2.15}$$

they extend smoothly to Σ^{n+1}. The same is true of the second derivatives $\partial^2 X^A/\partial x^\alpha \partial x^\beta$.

We deduce from these equations that the components $\partial_\alpha u = \frac{\partial X^A}{\partial x^\alpha}\partial_A u$ are linear with smooth coefficients on Σ^{n+1} in the components of $\partial_A u$. To compute the second derivatives $\tilde{\nabla}_\alpha \partial_\beta u$ in terms of $\nabla_A \partial_B u$ we use the relation between the Christoffel symbols of two conformal metrics and we obtain

$$\tilde{\nabla}_\alpha \partial_\beta \tilde{u} = \frac{\partial X^A}{\partial x^\alpha}\frac{\partial X^B}{\partial x^\beta}\{\nabla_A \partial_B(\Omega^{(n-1)/2}u) - \Omega^{-1}(\delta^C_A \partial_B \Omega + \delta^C_B \partial_A \Omega - \gamma^{CD}\gamma_{AB}\partial_D \Omega)\partial_C(\Omega^{(n-1)/2}u)\}. \tag{2.16}$$

Simple calculations give the following results.

1.

$$\frac{\partial X^A}{\partial t}\Omega^{-1}\partial_A \Omega = -\sin T \cos R, \qquad \frac{\partial X^A}{\partial r}\Omega^{-1}\partial_A \Omega = -\sin R \cos T. \tag{2.17}$$

Therefore the functions $\frac{\partial X^A}{\partial x^\alpha}\Omega^{-1}\partial_A \Omega$ extend to smooth functions on Σ^{n+1}.

2.

$$\gamma^{CD}\partial_C \partial_D \Omega = -\sin^2 T + \sin^2 R = \Omega(\cos t - \cos R), \tag{2.18}$$

Hence $\Omega^{-1}\gamma^{CD}\partial_C \partial_D \Omega$ extends to a smooth function on Σ^{n+1}. Remark that $\partial\Omega$ is a null vector on the boundary $\partial\mathcal{V}$.

3. The expression of Ω and the equalities

$$\frac{\partial T}{\partial t} = \frac{\partial R}{\partial r}, \quad \frac{\partial T}{\partial r} = \frac{\partial R}{\partial t} \tag{2.19}$$

imply that $\frac{\partial X^A}{\partial x^\alpha}\frac{\partial X^B}{\partial x^\beta}\Omega^{-1}\nabla^2_{AB}\Omega$ extends to a smooth function on Σ^{n+1}.

Using these results we see that Equation (2.9) on M^{n+1} is equivalent to a quasilinear wave equation on $\mathcal{V} \subset \Sigma^{n+1}$ of the form

$$\{\gamma^{AB} - \Omega^{n-1-\frac{3+n}{2}} m^{AB}(\Omega^{(n-1)/2}, u, \partial_A u)\}\nabla_B \partial_A u = \Omega^{n-1-\frac{3+n}{2}} k(\Omega^{(n-1)/2}, u, \partial_A u), \tag{2.20}$$

where m^{AB} and k are smooth functions of $\Omega^{(n-1)/2}$, u and $\partial_A u$, which vanish for $u = 0$. This equation extends to a quasilinear equation with continuous coefficients for u on Σ^{n+1} if

$$n - 1 - \frac{3+n}{2} = \frac{n-5}{2} \geq 0, \quad \text{i.e.} \quad n \geq 5. \tag{2.21}$$

The equation is hyperbolic as long as $\gamma^{AB} - \Omega^{\frac{n-5}{2}} m^{AB}$ is of Lorentzian signature, i.e. if u is small enough. It has smooth coefficients if in addition n is odd. □

The derivative of the function $\Omega^{1/2}$ is singular when $\Omega = 0$; when n is even sufficient regularity of the coefficients of the transformed equation, for the

applicability of the local existence theorem, holds only under some additional assumptions on the coefficients of the original equation[10].

2.3 *Cases $n = 3$, the null condition*

Recall (Definition XI.5.5) that Equation (2.6) satisfies the Christodoulou–Klainerman null condition if the quadratic form $F''(0,0,0)$, second derivative of F at $\tilde{u} = 0$ is a quadratic form which vanishes on tensors constructed through tensor products with null vectors ℓ of the Minkowski metric. In the case of the system (2.9) the null condition reduces to

$$h_{u^C}^{\alpha\beta\prime}(0)X^C\ell_\alpha\ell_\beta Z^A = 0, \quad Q_{BC}^{\alpha\beta,A}(0)\ell_\alpha Y^B\ell_\beta Y^C = 0 \tag{2.22}$$

for all X, Y, Z and vector ℓ null for the Minkowski metric.

2.3.1 *Case of one scalar equation*

When the unknown u is one scalar function the conditions (2.22) are equivalent to

$$h_u^{\alpha\beta\prime}(0)\ell_\alpha\ell_\beta = 0, \quad Q^{\alpha\beta}(0)\ell_\alpha\ell_\beta = 0 \tag{2.23}$$

for all vector ℓ null for the Minkowski metric

Theorem 2.2 *If the coefficients of Equation (2.9) depend smoothly on u and Du and satisfy the null condition (2.23), then the transformed equation on Σ^{n+1} has smooth coefficients for $n = 3$.*

Proof The null condition implies that, with $c^{\alpha\beta}$ and $d^{\alpha\beta}$ smooth in their arguments, $\tilde{u} \equiv \Omega^{(n-1)/2}u$

$$Q^{\alpha\beta}(\tilde{u}) = k\tilde{g}^{\alpha\beta} + \tilde{u}d^{\alpha\beta}(\Omega^{(n-1)/2}, u) \tag{2.24}$$

$$h^{\alpha\beta}(\tilde{u}) = k_1\tilde{g}^{\alpha\beta}\tilde{u} + \tilde{u}^2c^{\alpha\beta}(\Omega^{(n-1)/2}, u). \tag{2.25}$$

We use these equalities and the properties

$$\eta^{\alpha\beta}\partial_\alpha f_1\partial_\beta f_2 = \Omega^2\gamma^{AB}\partial_A f_1\partial_B f_2 \tag{2.26}$$

for any two functions, and in the case of Ω

$$\eta^{\alpha\beta}\partial_\alpha\Omega\partial_\beta\Omega = \Omega^2\gamma^{AB}\partial_A\Omega\partial_B\Omega = \Omega^3(\cos T - \cos\alpha) \tag{2.27}$$

to prove that the transformed equation extends smoothly to Σ^{n+1} if $n = 3$. □

[10] Satisfied in particular for wave maps into symmetric spaces. The results extend a theorem proved for $n = 2$ by Choquet-Bruhat, Y. and Gu Chao Hao (1989) *C. R. Acad. Sci. Paris*, serie 1, **308**, 167–70.

2.3.2 *Case of a system*

We consider the case where $\tilde{u} = (\tilde{u}^A)$, a set of scalar functions. The quadratic forms $\partial_\alpha u^A \partial_\beta u^B - \partial_\beta u^A \partial_\alpha u^B$ obviously vanish if one replaces the derivatives by a covariant vector ℓ. Sufficient condition for global existence for small data are as in the case of one equation $h'_{u^C}(0)$ proportional to the Minkowski metric, but now $\tilde{Q}(0)$ can be a linear combination of the forms

$$\eta^{\alpha\beta}\partial_\alpha \tilde{u}^A \partial_\beta \tilde{u}^B \quad \text{and} \quad \partial_\alpha u^A \partial_\beta u^B - \partial_\beta u^A \partial_\alpha u^B,$$

called **standard null forms**.

2.4 *Wave maps*

Wave maps are semilinear second-order differential equations on a Lorentzian manifold, which take their values in a pseudo-Riemannian manifold. The Penrose map transforms wave maps from the Minkowski spacetime into wave maps from the open set $\mathcal{V}$ of the Einstein cylinder. The method used to prove global existence for small data for a quasilinear quasidiagonal systems of wave equations applies to wave maps.

It is trivial to check that wave maps satisfy the null condition: they obey a global existence theorem near a constant map[11] for $n = 3$. It has been proved[12] that for symmetric space targets the theorem holds also for $n = 2$.

3 H. Friedrich conformal system, $n + 1 = 4$

As we already pointed out, the Einstein equations, like wave equations with a mass term, are not conformally invariant. However, on a four-dimensional spacetime, the Weyl tensor which is considered to represent, in some sense, the radiation part of the gravitational field, possesses remarkable properties under conformal transformations of the spacetime metric.

The Einstein–Bianchi equations on a four-dimensional spacetime (see Chapter 8) lead to local solutions of the Cauchy problem, when coupled with equations determining connection and metric from the Riemann tensor. H. Friedrich has introduced, using the Newman–Penrose formalism, a symmetric hyperbolic first-order system for the Weyl tensor which he couples with an equation for the conformal factor, and the Einstein equations for the conformal metric.

3.1 *Equations*

A **spin-2 field** is a 4-tensor possessing the symmetries and antisymmetries of the Riemann tensor, and, in addition, a vanishing trace. The Weyl tensor is a

[11] Choquet-Bruhat, Y. (1986) *C. R. Acad. Sci. Paris* serie I, **303**(4), 109–11 and (1987) *Ann. Inst. Poincaré*, **46**(1), 97–111. Sideris showed, more generally, global existence near a geodesic map.

[12] Choquet-Bruhat and Gu Chao Hao (1989) *loc. cit.*

spin-2 field. The following theorem gives an identity under a conformal change of spacetime metric.

Theorem 3.1 *Let $\tilde{g}$ and g be two conformal metrics on a four-dimensional manifold V*

$$\tilde{g}_{\alpha\beta} = \omega^{-2} g_{\alpha\beta}.$$

1. *(Penrose, 1965) Let $W_\alpha{}^\beta{}_{\lambda\mu}$ be a spin-2 field, $W_\alpha{}^\beta{}_{\lambda\beta} = 0$, then the following identity holds, with ∇ and $\tilde{\nabla}$ respectively the covariant derivatives in the metrics g and $\tilde{g}$,*

$$\sum_{\gamma,\lambda,\mu} \tilde{\nabla}_\gamma W_\alpha{}^\beta{}_{\lambda\mu} \equiv \omega \sum_{\gamma,\lambda,\mu} \nabla_\gamma (\omega^{-1} W_\alpha{}^\beta{}_{\lambda\mu}). \tag{3.1}$$

2. *(Friedrich 1981) Define the traceless 2-tensor Σ by*

$$\Sigma_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{4} g_{\alpha\beta} R. \tag{3.2}$$

The tensors Σ and $\tilde{\Sigma}$ of the conformal metrics g and $\tilde{g}$ are linked by the identities

$$\omega \tilde{\Sigma}_{\alpha\beta} \equiv \omega \Sigma_{\alpha\beta} + 2\nabla_\alpha \partial_\beta \omega - \frac{1}{2} g_{\alpha\beta} \nabla^\lambda \partial_\lambda \omega, \tag{3.3}$$

and

$$2\tilde{\Sigma}_{\alpha\beta}\partial^\beta \omega + \frac{1}{6}\omega^{-1}\partial_\alpha \tilde{R} \equiv 2\Sigma_{\alpha\beta}\partial^\beta \omega + \frac{1}{6}\omega \partial_\alpha R + \frac{1}{3} R \partial_\alpha \omega + \nabla_\alpha \nabla^\lambda \partial_\lambda \omega. \tag{3.4}$$

Proof Straightforward computations using the identities in Appendix VI. □

Friedrich uses the identities deduced from the Bianchi identities (see Chapter 8) satisfied respectively by the Weyl tensor of the metric g and the same Weyl tensor of the metric $\tilde{g}$ which is solution of the vacuum Einstein equations, $\tilde{R}_{\alpha\beta} = 0$,

$$\nabla_\alpha W_\beta{}^\alpha{}_{,\lambda\mu} \equiv -\frac{1}{2}(\nabla_\lambda \Sigma_{\beta\mu} - \nabla_\mu \Sigma_{\beta\lambda}) \quad \text{and} \quad \tilde{\nabla}_\alpha W_\beta{}^\alpha{}_{,\lambda\mu} = 0;$$

he obtains equations linking the tensors Σ and $d := \omega^{-1} W$,

$$\frac{1}{2}(\nabla_\lambda \Sigma_{\beta\mu} - \nabla_\mu \Sigma_{\beta\lambda}) + \partial_\alpha \omega d_\beta{}^\alpha{}_{,\lambda\mu} + \frac{1}{24}(g_{\beta\mu}\partial_\lambda R - g_{\beta\lambda}\partial_\mu R) = 0.$$

Friedrich remarks that the scalar curvature R can take arbitrary values by conformal rescaling of the metric. He considers it as a given gauge variable, and replaces it in the equations by a given function r. He obtains the following theorem.

Theorem 3.2 *(**Friedrich conformal system**) Let $\tilde{g}$ be a metric solution of the vacuum Einstein equations, possibly with a cosmological constant Λ,*

$$\tilde{R}_{\alpha\beta} - \Lambda \tilde{g}_{\alpha\beta} = 0, \quad \textit{hence} \quad \tilde{R} = 4\Lambda \quad \textit{and} \quad \tilde{\Sigma}_{\alpha\beta} = 0.$$

Then the conformal metric

$$g := \omega^2 \tilde{g},$$

the scalar function ω and the tensors Σ, d, with their symmetries and traceless properties, satisfy the differential system

$$R_{\alpha\beta} = \Sigma_{\alpha\beta} + \frac{1}{4} g_{\alpha\beta} r, \tag{3.5}$$

$$\omega \Sigma_{\alpha\beta} + 2\nabla_\alpha \partial_\beta \omega - \frac{1}{2} g_{\alpha\beta} \nabla^\lambda \partial_\lambda \omega = 0, \tag{3.6}$$

$$2\Sigma_{\alpha\beta} \partial^\beta \omega + \frac{1}{6} \omega \partial_\alpha r + \frac{1}{3} r \partial_\alpha \omega + \nabla_\alpha \nabla^\lambda \partial_\lambda \omega = 0, \tag{3.7}$$

$$\sum_{\gamma,\lambda,\mu} \nabla_\gamma d_\alpha{}^\beta{}_{\lambda\mu} = 0, \tag{3.8}$$

$$\frac{1}{2}(\nabla_\lambda \Sigma_{\beta\mu} - \nabla_\mu \Sigma_{\beta\lambda}) + \partial_\alpha \omega d_\beta{}^\alpha{}_{\lambda\mu} + \frac{1}{24}(g_{\beta\mu} \partial_\lambda r - g_{\beta\lambda} \partial_\mu r) = 0. \tag{3.9}$$

3.2 *Friedrich hyperbolic system*

Friedrich uses a wave gauge for the metric g. He introduces as an auxiliary unkown the function $\sigma := \nabla^\lambda \partial_\lambda \omega$; taking the covariant divergence of (3.3) and using (3.2), with the tracelessness of $\Sigma_{\alpha\beta}$, replaces these equations for ω by the well-posed system of linear wave equations for ω and σ when $\Sigma_{\alpha\beta}$ and r are known (one uses the equality $\nabla^\alpha \Sigma_{\alpha\beta} = \partial_\beta r$ coming form the definitions and the Bianchi identity)

$$\nabla^\alpha \partial_\alpha \omega = \sigma,$$

$$\nabla^\alpha \partial_\alpha \sigma - \omega \Sigma^{\alpha\beta} \Sigma_{\alpha\beta} + 2\partial_\beta r \partial^\beta \omega + \frac{1}{2} \partial^\alpha \omega \partial_\alpha r + \frac{1}{6} \omega \nabla^\alpha \partial_\alpha r + \frac{1}{3} r\sigma = 0.$$

Friedrich uses a Newman–Penrose representation for the tensor d to construct a symmetric hyperbolic first-order system satisfied by his unknowns and obtain a metric g and a conformal factor ω with, for instance, initial data near to initial data for the Einstein cylinder and the Penrose conformal factor Ω. The solution will exist on a slice of the Einstein cylinder up to any given time T if these initial data are small enough. However, the metric $\tilde{g}$ satisfies the original Einstein equations only if it also satisfies the Einstein constraints. There is, as pointed out in the case of the global solution of Yang–Mills equations by the Penrose transform, a problem with the fall-off at space infinity of the original data necessary for the smoothness of the transformed data required for the existence of a (even

local) solution of the transformed evolution hyperbolic system. Friedrich circumvents this difficulty by considering original data on an hyperboloid which turns null at space infinity. Large classes of admissible such solutions of constraints have been constructed by Andersson, Chruściel and Friedrich[13]. In recent papers Friedrich introduces more elaborate formalisms and sophisticated structure of conformal space infinity[14].

It is also possible to write, without using the Newman–Penrose formalism, a Leray hyperbolic system[15] satisfied by Friedrich's unknowns.

Remark 3.3 *If g is smooth, $\tilde{g}$ becomes singular for $\omega = 0$.*

4 Einstein's equations in higher dimensions

Einstein's equations in wave coordinates are on R^{n+1} a set quasidiagonal quasilinear wave equations for the components $g_{\mu\nu}$ of the spacetime metric g of the form considered in Section 2:

$$g^{\alpha\beta}\frac{\partial^2 g_{\mu\nu}}{\partial x^\alpha \partial x^\beta} + Q_{\mu\nu}(g)\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial x}\right). \tag{4.1}$$

The Penrose transform gives the global existence of a solution of the Cauchy problem for small (i.e. here near Minkowskan) initial data when[16] $n \geq 5$. There is however a difficulty for application to the Einstein equations, as we already mentioned, due to the fall-off at infinity required from the original initial data for the local existence theorem to be applicable to the transformed equations. Indeed, the positive mass theorem, a consequence of the Einsteinian constraints, implies that such a fall-off holds only for the Minkowski spacetime itself. We use[17] another conformal mapping which circumvents this problem.

4.1 *Conformal mapping*

To prove our global existence result we use a mapping $\phi : x \mapsto y$ from the future timelike cone with vertex 0, $I^+_{\eta,x}$, of a Minkowski spacetime which we denote (R^{n+1}_x, η_x) into the past timelike cone, $I^-_{\eta,y}$ with vertex $y = \phi(x)$ of another Minkowski spacetime, (R^{n+1}_y, η_y). This map is defined by

$$\phi : \quad I^+_{\eta,x} \to R^{n+1}_y \quad \text{by} \quad x^\alpha \mapsto y^\alpha := \frac{x^\alpha}{\eta_{\lambda\mu}x^\lambda x^\mu}\ . \tag{4.2}$$

[13] Andersson, L. and Chruściel, P. (1994), *Commun. Math. Phys.*, **161**(3), 533–68.

[14] Friedrich, H. (2004) in *The Einstein Equations and the Large Scale Behavior of Gravitational Fields* (eds. P. Chruściel and H. Friedrich), Birkhäuser, and references therein.

[15] Choquet-Bruhat, Y. and Novello, M. (1987) *C. R. Acad. Sci. Paris*, **305**, 155–60; see e.g. CB-DM2 V 7.

[16] The $3+1$ Einstein equations in harmonic coordinates do not satisfy the null condition; see Section 15.7.

[17] Choquet-Bruhat, Y., Chruściel, P., and Loiselet, J. (2006) *loc. cit.* See also a previous paper of Anderson, M. and Chruściel, P. (2005) *Commun. Math. Phys.*, **260**, 557–77, which used another method.

Lemma 4.1 *The mapping ϕ is a confomorphism from $I^+_{\eta,x}$ onto $I^-_{\eta,y}$.*

Proof The map ϕ is differentiable and admits the inverse

$$\phi^{-1} : y \mapsto x \quad \text{by} \quad x^\alpha := \frac{y^\alpha}{\eta_{\lambda\mu} y^\lambda y^\mu}. \tag{4.3}$$

An easy computation gives a conformal relation under the mapping ϕ between the Minkowski metrics of R^{n+1}_x and R^{n+1}_y. It holds that, with η the diagonal quadratic form $(-1, 1, \ldots, 1)$,

$$\eta_{\alpha\beta} dx^\alpha dx^\beta = \Omega^{-2} \eta_{\alpha\beta} dy^\alpha dy^\beta \tag{4.4}$$

where Ω is a function on $I^-_{\eta,y}$ given by

$$\Omega := \eta_{\alpha\beta} y^\alpha y^\beta. \tag{4.5}$$

□

4.2 *Transformed equations*

We set as before

$$g_{\mu\nu} := \eta_{\mu\nu} + h_{\mu\nu}. \tag{4.6}$$

We consider $h := (h_{\mu\nu})$ as a set of scalar functions on R^{n+1}_x. Einstein's equations in wave coordinates, for the unknowns $h := (h_{\mu\nu})$, are then

$$\eta^{\alpha\beta} \frac{\partial^2 g_{\mu\nu}}{\partial x^\alpha \partial x^\beta} - g^{\alpha\beta} \frac{\partial^2 h_{\lambda\mu}}{\partial x^\alpha \partial x^\beta} + Q_{\mu\nu}(\eta + h) \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial x} \right)^2 = 0 \tag{4.7}$$

The relation between the wave operator on scalar functions in two conformal metrics, here both on flat spaces, gives the identity

$$\eta^{\alpha\beta} \frac{\partial^2 (\Omega^{-\frac{n-1}{2}} h \circ \phi^{-1})}{\partial y^\alpha \partial y^\beta} \equiv \Omega^{-\frac{n+3}{2}} \left(\eta^{\alpha\beta} \frac{\partial^2 h}{\partial x^\alpha \partial x^\beta} \right) \circ \phi^{-1}. \tag{4.8}$$

The system (4.7) on R^{n+1}_x is therefore equivalent to the following system on R^{n+1}_y

$$\eta^{\alpha\beta} \frac{\partial^2 \hat{h}}{\partial y^\alpha \partial y^\beta} + \Omega^{-\frac{n+3}{2}} \left\{ (g^{\alpha\beta} - \eta^{\alpha\beta}) \frac{\partial^2 h}{\partial x^\alpha \partial x^\beta} + Q(\eta + h) \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial x} \right) \right\} \circ \phi^{-1} = 0 \tag{4.9}$$

for the set of scalar functions

$$\hat{h}_{\lambda\mu} := \Omega^{-\frac{n-1}{2}} h_{\lambda\mu} \circ \phi^{-1}.$$

We note that $g^{\alpha\beta} - \eta^{\alpha\beta}$ is a rational function of $h := (h_{\mu\nu})$ with numerator a linear function of h and with denominator bounded away from zero as long as $g_{\alpha\beta}$ is non-degenerate. Therefore

$$(g^{\alpha\beta} - \eta^{\alpha\beta}) \circ \phi^{-1} \equiv \Omega^{\frac{n-1}{2}} f^{\alpha\beta}(\Omega^{\frac{n-1}{2}}, \hat{h})$$

where $f^{\alpha\beta}$ is a rational function of $\hat{h}$ and $\Omega^{\frac{n-1}{2}}$ with denominator bounded away from zero as long as $g_{\alpha\beta}$ is non-degenerate.

Theorem 4.2 *The bijection ϕ transforms the Einstein equations in wave coordinates on $I^+_{\eta,x}(0)$ into a quasidiagonal, quasilinear system on $I^-_{\eta,y}(0)$ which extends smoothly to R^{n+1}_y, as long as the metric $\eta + \Omega^{\frac{n-1}{2}}\hat{h}$, is non-degenerate, n is odd and $n \geq 5$.*

Proof We use the definitions of Ω and of the maping ϕ to compute the various coefficients of Equation (4.9). We have

$$\frac{\partial h}{\partial x^\lambda} \circ \phi^{-1} \equiv A^\alpha_\lambda \frac{\partial (h \circ \phi^{-1})}{\partial y^\alpha}, \tag{4.10}$$

with

$$A^\alpha_\mu := \frac{\partial y^\alpha}{\partial x^\mu} \equiv \Omega \delta^\alpha_\mu - 2y^\alpha y_\mu \quad \text{and} \quad \frac{\partial \Omega}{\partial y^\alpha} = 2\eta_{\alpha\beta} y^\beta := 2y_\alpha.$$

We see that these quantities are bounded on any bounded set of R^{n+1}_y.

Elementary but somewhat lengthy computations[18] then give

$$\frac{\partial^2 h}{\partial x^\lambda \partial x^\mu} \circ \phi^{-1} = B^{\alpha\beta}_{\lambda\mu} \frac{\partial^2 (h \circ \phi^{-1})}{\partial y^\alpha \partial y^\beta} + C^\alpha_{\lambda\mu} \frac{\partial (h \circ \phi^{-1})}{\partial y^\alpha} \tag{4.11}$$

where the coefficients B and C, given by.

$$B^{\alpha\beta}_{\lambda\mu} \equiv \Omega^2 \delta^\beta_\lambda \delta^\alpha_\mu - 2\Omega(y^\beta y_\lambda \delta^\alpha_\mu + y^\alpha y_\mu \delta^\beta_\lambda) + 4y^\alpha y^\beta y_\lambda y_\mu \tag{4.12}$$

and

$$C^\alpha_{\lambda\mu} \equiv -2\Omega(y_\lambda \delta^\alpha_\mu + y_\mu \delta^\alpha_\lambda + y^\alpha \eta_{\lambda\mu}) + 8y^\alpha y_\lambda y_\mu), \tag{4.13}$$

are bounded on any bounded subset of R^{n+1}_y. Replacing $h \circ \phi^{-1}$ by $\Omega^{\frac{n-1}{2}}\hat{h}$ gives after some more elementary computations

$$\frac{\partial (h \circ \phi^{-1})}{\partial y^\alpha} \equiv \frac{\partial (\Omega^{\frac{n-1}{2}} \hat{h})}{\partial y^\alpha} = \Omega^{\frac{n-1}{2}} \frac{\partial \hat{h}}{\partial y^\alpha} + (n-1)\Omega^{\frac{n-3}{2}} y_\alpha \hat{h} \tag{4.14}$$

and

$$\begin{aligned} \frac{\partial^2 (h \circ \phi^{-1})}{\partial y^\alpha \partial y^\beta} \equiv{}& \Omega^{\frac{n-1}{2}} \frac{\partial^2 \hat{f}}{\partial y^\alpha \partial y^\beta} + (n-1)\Omega^{\frac{n-3}{2}} \\ &\times \left(y_\beta \frac{\partial \hat{f}}{\partial y^\alpha} + y_\alpha \frac{\partial \hat{f}}{\partial y^\beta} \right) + \frac{n-1}{2} \Omega^{\frac{n-5}{2}} D_{\alpha\beta} \hat{f} \end{aligned} \tag{4.15}$$

with

$$D_{a\beta} := 2(n-3) y_\alpha y_\beta + 2\eta_{\alpha\beta}\Omega. \tag{4.16}$$

[18] See details in Choquet-Bruhat, Chruściel, and Loiselet (2006).

Reassembling these results and using the properties

$$\eta^{\alpha\beta}\frac{\partial\Omega}{\partial y^\beta}\frac{\partial\Omega}{\partial y^\alpha}\equiv 4\Omega,\quad \text{and}\quad \eta^{\alpha\beta}\frac{\partial^2\Omega}{\partial y^\alpha\partial y^\beta}\equiv 2(n+1), \tag{4.17}$$

we see that if $n\geq 5$ and n is odd, the system (4.9) is a smooth system on R_y^{n+1} of quasidiagonal quasilinear second-order hperbolic differential equations, as long as the metric $\eta+\Omega^{\frac{n-1}{2}}\hat{h}$ is non-degenerate. □

4.3 *Local Cauchy problem in R_x^{n+1}, $n\geq 5$ and odd*

4.3.1 *Initial data*

We consider the Cauchy problem on R_x^{n+1} for the Einstein equations in wave coordinates. We assume that the initial data $(\bar{g},K)$ are given on a submanifold $M_x:=\{x^0=2\lambda\}$ of R_x^{n+1} and satisfy the Einstein constraints. We choose the initial lapse and shift and their first time derivatives so that the harmonicity conditions are everywhere initially zero. Any globally hyperbolic solution $(\mathcal{V}_x\subset R_x^{n+1},g)$ of the Einstein equation in wave coordinates taking these initial data is then a solution of the full Einstein equations. Such a solution exists (see Chapter 6) if the initial data belongs to $H_{s+1}^{u.loc}\times H_s^{u.loc}$, $s>\frac{n}{2}$.

For the global existence theorem one supposes that the initial data, for the Einstein equations in R_x^{n+1} in wave coordinates, coincide with the data for the Schwarzschild metric in wave coordinates outside a ball B_{R_x} of radius $r=R_x$. Large families of such initial data sets can be constructed[19], such that they are arbitrarily close to those for Minkowski spacetime.

By the general uniqueness theorem for quasilinear wave equations, the solution $(\mathcal{V}_x,g)$ coincides with the Schwarzschild spacetime, in wave coordinates, in the domain of dependence of the Schwarzschild region, $M_x\backslash B_{R_x}$.

4.3.2 *Domain of dependence of the Schwarzschild initial data*

The Schwarzschild metric $g_{Schw,m}$ with mass parameter m, in any dimension $n\geq 3$, is in standard coordinates

$$g_{Schw,m}=-\left(1-\frac{2m}{r^{n-2}}\right)dt^2+\frac{dr^2}{1-\frac{2m}{r^{n-2}}}+r^2d\Omega^2 \tag{4.18}$$

where $d\Omega^2$ is the canonical metric on S^{n-1}.

We have shown in Chapter 4[20] that there are wave coordinates x^α where for the Schwarzschild metric in space dimension $n>4$ takes the form

$$(g_{Schw,m})_{\mu\nu}=\eta_{\mu\nu}+\frac{1}{r^{n-2}}f_{\mu\nu}\left(m,\frac{1}{r},\frac{\bar{x}}{r}\right),\quad \bar{x}:=(x^i), \tag{4.19}$$

with $f_{\mu\nu,m}$ analytic functions of their arguments for large enough r and vanishing for $m=0$.

[19] Chruściel, P. and Delay, E., Corvino, J.; see references in last section of Chapter 7.
[20] See the last section.

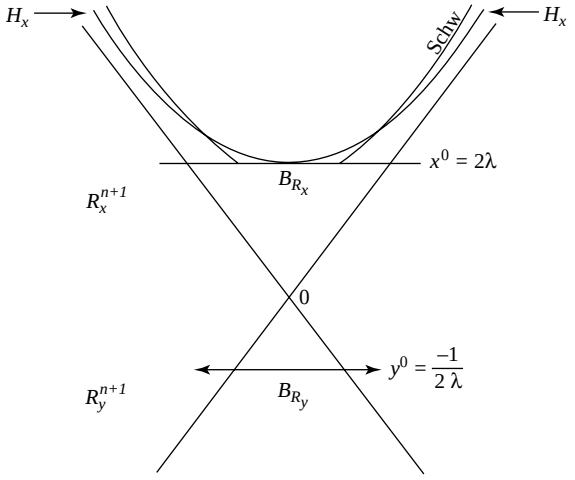


Figure 15.1

The boundary of the future domain of dependence, $\partial\mathcal{D}_x^+(M_x\backslash B_{R_x})$ is threaded by null radial outgoing geodesics of the Schwarzschild metric issued from the boundary ∂B_{R_x}, which are solutions of the differential equation

$$\frac{dt}{dr} = \left(\frac{g_{m,rr}}{g_{m,tt}}\right)^{\frac{1}{2}} \tag{4.20}$$

such that $t(R_x) = 2\lambda$, i.e.

$$x^0 \equiv t = 2\lambda + \int_R^r \left(\frac{g_{m,rr}}{g_{Sm,tt}}\right)^{\frac{1}{2}} dr. \tag{4.21}$$

Using the bounds found for the Schwarzschild metric if $n \geq 5$, with C some constant

$$\left(\frac{g_{Sch,rr}}{g_{Sch,tt}}\right)^{\frac{1}{2}} - 1 \leq C\frac{m}{r^{n-2}}, \tag{4.22}$$

we deduce from (4.21) that on the boundary $\partial\mathcal{D}_x^+(M_x\backslash B_{R_x})$ we have

$$C_1 \geq t - r \geq C_2, \tag{4.23}$$

with

$$C_1 := 2\lambda - R + Cm\frac{1}{(n-3)R_x^{n-3}}, \qquad C_2 := 2\lambda - R_x - Cm\frac{1}{(n-3)R_x^{n-3}}. \tag{4.24}$$

We see that if we choose $\lambda > R_x$ and if m is small enough then $C_2 > 0$ hence $\partial\mathcal{D}^+(M_x\backslash B_{R_x})$ is interior to $I_{\eta,x}^+(0)$.

4.4 *Global Cauchy problem*

To solve the global Cauchy problem on R_x^{n+1} with data on $x^0 = 2\lambda$ for the original system (4.7), we consider the local Cauchy problem for the transformed system on R_y^{n+1} (4.9) with initial data in $H_{s+1} \times H_s$, $s > \frac{n}{2}$, on a ball $B_{R_y} \subset M_y := \{y^0 = -\frac{1}{2\lambda}\}$ of radius $R_y > \frac{1}{2\lambda}$.

The hyperplane $y^0 = -\frac{1}{2\lambda}$ is the image by ϕ^{-1} of the hyperboloid

$$H_x := \{x^0 = \lambda + \sqrt{\lambda^2 + r^2}\}.$$

To show that the considered Cauchy data for the transformed system can be determined, we prove the following lemma.

Lemma 4.3 *If the original initial data, including m, are small enough, the hyperboloid H_x is included in the domain $\mathcal{V}_x$ of the existence of the local solution on R_x^{n+1}.*

Proof The hyperboloid H_x is the union of two subsets

$$S_1 := H_x \cap \{x^0 - r\} \geq a \quad \text{and} \quad S_2 := H_x \cap \{x^0 - r\} \leq a \tag{4.25}$$

The subset S_2 is included in the Schwarzschild spacetime region. On the subset S_1 it holds that

$$\lambda + \sqrt{r^2 + \lambda^2} \geq r + a \tag{4.26}$$

A simple computation shows that r is bounded on S_1 if $\lambda > R_x$ and m is small, because, using the value of a, one finds that:

$$r \leq \frac{R_x(2\lambda - R_x) - O(m)}{2(\lambda - R_x)}; \tag{4.27}$$

then x^0 is also bounded on S_1. This subset is therefore included in the existence domain $\mathcal{V}_x$ of the local solution with Cauchy data on M_x, for small enough Cauchy data. □

We deduce the initial data on M_y by the mapping ϕ from the values on H_x of the local solution in R_x^{n+1} and its first derivative.

The image initial data $\hat{h}_2$ for $\hat{h}$ is deduced from the Schwarzschild metric in wave coordinates, which is static; we have

$$\hat{h}_2(\bar{y}) = [\Omega^{-\frac{n-1}{2}}(f_m \circ \phi^{-1})]\left(y^0 = -\frac{1}{2\lambda}, \bar{y}\right), \tag{4.28}$$

We see that $\hat{h}_2$ is a smooth function of y, except at the origin $\bar{y} = 0$, if $n \geq 3$. A simple calculation shows that the same is true of $\partial\hat{h}/\partial y^0$ on $\phi(H_x \cap S_2)$ as soon as $n \geq 5$.

On $\phi(H_x \cap S_1)$, the initial data $\hat{h}_1$ are deduced from the values on the uniformly spacelike submanifold $H_x \cap S_1$ of the solution in $\mathcal{V}_x$ and its derivative,

by the restriction of ϕ to a neighbourhood of $H_x \cap S_1$, where ϕ is a smooth diffeomorphism.

We deduce from these results that, for small enough Cauchy data (including small m in the Schwarzschild part) on M_x we can compute Cauchy data with required properties on M_y for the transformed system.

The corresponding Cauchy problem admits a solution $\hat{h}$, with $\eta + \hat{h}$ a non-degenerate Lorentzian metric, in a neighbourhood of

$$\mathcal{D}_y := I^-_{\eta,y}(0) \cap \left\{ y^0 \geq -\frac{1}{2\lambda} \right\}, \tag{4.29}$$

if the initial data are small enough. The Einstein equations in wave coordinates on R^{n+1}_x have then a solution in $\mathcal{D}_x := \phi^{-1}(\mathcal{D}_y)$, the whole future of the hyperboloid H_x, i.e. finally in the whole future of the initial manifold $x^0 = 2\lambda$.

4.5 *Conclusion*

We have proved the following theorem.

Theorem 4.4 *Let n be odd and $n \geq 5$. Let there be given on R^n gravitational data, a perturbation of the Minkowski data given by sets of functions $\bar{h}_{\mu\nu}$ and $\overline{\frac{\partial}{\partial x^0} h_{\mu\nu}}$. Suppose that these data satisfy the Einstein constraints and the initial harmonicity conditions and coincide with the Schwarzschild data in wave coordinates outside a ball B of finite Euclidean radius. Then, if these functions are small enough in $H_{s+1} \times H_s$ norm on B, and m is small, the data admit a complete Einsteinian development (R^{n+1}, g).*

5 Christodoulou–Klainerman theorem

The proof of the non-linear stability of Minkowski spacetime M^4 by Christodoulou and Klainerman relies on generalized energy estimates for the Weyl tensor linked with its conformal properties, and on relations between the connection, the metric and the Weyl tensor, together with a bootstrap argument. It is impossible to condense in a few pages the 500 pages of the proof, and impossible to state, without introducing a wealth of notations and auxiliary lemmas, all the properties that the authors obtain along the way. We will state the CK theorem in the least refined forms given by the authors; more complete formulations can be found in their book. We will only give some hints on the methods used to prove their results.

5.1 *CK main theorem*

We give some definitions necessary to understand the formulation of the theorem[21].

[21] We change some of the CK notations for notations used in preceding chapters. One says that f$\in o_m(r^k)$ .if $\partial^l \mathrm{f} \in o(r^{k-l})$, $0 \leq l \leq m$.

Definition 5.1 *An initial data set $(M, \bar{g}, K)$ on a manifold M Euclidean at infinity is called* ***strongly asymptotically Euclidean*** *if it is complete and if, as $r := \{\sum (x^i)^2\}^{\frac{1}{2}} \to \infty$,*

$$\bar{g}_{ij} = \left(1 + \frac{2m}{r}\right)\delta_{ij} + o_4(r^{-\frac{3}{2}}) \qquad K_{ij} = o_3(r^{-\frac{5}{2}}).$$

With this definition the ADM mass is m, the linear momentum is zero, and the angular momentum is well defined. One says that the initial data are in the centre of mass frame.

Instead of confronting $\bar{g}_0$ directly with the Euclidean metric, CK characterize as follows the smallness of an initial data set on a manifold Euclidean at infinity, with d_0 the distance to some point $x_0 \in M$ and $\sigma_0 := \{1 + d_0^2\}^{\frac{1}{2}}$.

Definition 5.2 *The smallness of a strongly asymptotically Euclidean initial data $(\bar{g}_0, K_0)$ set is measured by the infimum for $x_0 \in M$ of the following quantity*

$$Q(x_0; \bar{g}_0, K_0) := \underset{M}{Sup}(d_0^2 + 1)^3 |Ricci(\bar{g}_0)|^2 \tag{5.1}$$
$$+ \int_M \left\{ \sum_{l=0}^{3} (d_0^2 + 1)^{l+1} |\bar{\nabla}^l K_0|^2 + \sum_{l=0}^{1} (d_0^2 + 1)^1 |\bar{\nabla}^l B_0|^2 \right\} \mu_{\bar{g}_0}$$

with B the symmetric traceless ***Bach tensor***

$$B_{ij} := \frac{1}{2}(\eta_i{}^{hk}\bar{\nabla}_h \tilde{R}_{jk} + \eta_j{}^{hk}\bar{\nabla}_h \tilde{R}_{ik}), \quad \tilde{R}_{ij} = \bar{R}_{ij} - \frac{1}{3}\bar{g}_{ij}\bar{R}. \tag{5.2}$$

Note that, in dimension 3, the Bach tensor is the dual of the **Coton tensor**[22] which vanishes for conformally flat metrics.

We state the main theorem of Christodoulou and Klainerman, which they call the "second version".

Theorem 5.3 *There exists $\varepsilon > 0$ such that any strongly asymptotically Euclidean maximal initial data set which satisfies the Einstein vacuum constraints and satisfies the global smallness assumption*

$$\underset{x_0 \in M}{Inf}\, Q(x_0; \bar{g}_0, K_0) < \varepsilon \tag{5.3}$$

admits a globally hyperbolic vacuum Einsteinian development which is geodesically complete, regularly sliced by maximal submanifolds, with zero shift. Moreover, this development is globally asymptotically flat.

We give as a corollary what CK call the "first version of the main theorem". It results from the fact that Lorentzian manifolds near the Minkowski spacetime always admit a maximal hypersurface[23].

[22] See Appendix VI.

[23] Choquet-Bruhat, Y. (1976) *Ann. Scuola Norm. Pisa*, Serie IV No. 3, 361–78, in honour of J. Leray. The property has been shown to hold, more generally, in Lorentzian manifolds satisfying some uniform bounds (Bartnik, R. (1984) *Commun. Math. Phys.*, **94**, 155–75).

Corollary 5.4 *The hypothesis that the initial data set satisfies $tr_{\bar{g}_0} K = 0$ can be removed from the hypotheses.*

5.2 *Local existence*

CK work in maximal time gauge and zero shift. The constraint equations on a 3-manifold M read then

$$\bar{\nabla}_i \bar{K} ij = 0, \quad tr_{\bar{g}} \bar{K} = 0, \quad \bar{R} - |K|^2 = 0;$$

and their solution by the conformal method is straightforward (see Chapter 7).

The $3+1$ evolution system is

$$\partial_t g_{ij} = -2NK_{ij}, \quad \partial_t K_{ij} = N(\bar{R}_{ij} - 2K_{ih}K^h_j) - \bar{\nabla}_i \partial_j N.$$

We have shown in Chapter 8 that a solution of the vacuum Einstein equations for a metric with given shift (here $\beta = 0$) and given mean curvature $Tr_{\bar{g}} K = H$ of time slices (here $H = 0$) satisfies the hyperbolic system (remark that in dimension 3 the Riemann tensor is a linear function of the Ricci tensor)

$$\partial_t g_{ij} = -2NK_{ij}, \tag{5.4}$$

$$-\partial_t(N^{-1}\partial_t K_{ij}) + \bar{\nabla}_h \bar{\nabla}^h (NK_{ij}) - \partial_t(N^{-1}\bar{\nabla}_j \partial_i N) - 2\partial_t K_{im} K^m_j \tag{5.5}$$

$$-\bar{\nabla}_{(i}(K_{j)h}\partial^h N) - 2N\bar{R}^h_{\ ijm}K^m_h - N\bar{R}_{m(i}K^m_{j)} = 0.$$

This system is to be coupled with an elliptic equation for the lapse N which reads when $H = 0$:

$$\bar{\nabla}^i \partial_i N - |K|^2 N = 0, \quad |K|^2 \equiv K_{ij}K^{ij}. \tag{5.6}$$

We have shown that a solution of this system with initial data satisfying the constraints satisfy the original vacuum Einstein equations.

CK utilize the above system to prove a local existence theorem, with an appropriate fall-off at infinity of the solution. We state the theorem essentially in their own terms. The spaces $H_{s,\delta}$ are the weighted Sobolev spaces on $(M, \bar{g}_0)$ (see Appendix I).

Theorem 5.5 *(local existence) Let $(M, \bar{g}_0, K_0)$ be an initial data set, with $tr_{\bar{g}_0} K_0 = 0$, satisfying the vacuum constraints, and such that*

1. *$(M, \bar{g}_0)$ is a complete Riemannian manifold, with M diffeomorphic to R^3.*
2. *The isoperimetric constant $I(M, \bar{g}_0)$ is finite, where I is defined to be*

$$\underset{S}{Sup} \frac{V(S)}{A(S)^{\frac{3}{2}}}$$

   with S an arbitrary surface in M, $A(S)$ its area and $V(S)$ its volume.
3. *The Ricci curvature of $\bar{g}_0$ and the symmetric 2-tensor K_0 are such that*

$$Ricci(\bar{g}_0) \in H_{2,1} \quad K_0 \in H_{3,1}.$$

Then there exists a local in time, unique, vacuum future Einsteinian development, $(M \times [0, t_*), g)$, with M_t maximal slices and with zero shift, and $\bar{g}$, K such that

$$\bar{g} - \bar{g}_0 \in C^1([0, t_*)), H_{3,1}), \quad K \in C^0([0, t_*)), H_{3,1}),$$

and, furthermore

$$Ricci(\bar{g}) \in C^0([0, t_*)), H_{2,1}).$$

On the other hand if 3 is strengthened to

$$\sigma_0^3 |Ricci(\bar{g}_0)| \quad \textit{is bounded}, \quad B(\bar{g}_0) \in H_{1,3},$$

where B is the Bach tensor, then the solution satisfies in addition

$$\sigma_0^3 |Ricci(\bar{g})| \quad \textit{is bounded}, \quad B(\bar{g}) \in H_{1,3}.$$

Proof The proof is by iteration, solving successively the elliptic lapse equation for known $\bar{g}$ and K, and the hyperbolic system for $\bar{g}$ and K with a given N. The proof uses some global Sobolev inequalities satisfied by a function f on $(M, \bar{g})$, in particular, with I the isoperimetric bound

$$||f||_{L^{\frac{3}{2}}} \leq I^{\frac{2}{3}} ||\bar{\nabla} f||.$$

Energy type estimates applied to the wave equation satisfied by K lead to bounds in $L^2(M, g_0)$ for $\sigma_0 \bar{\nabla} K$ and $\sigma_0 \partial_t K$. CK indicate how a better fall-off at infinity is obtained by derivating from this equation other wave equations satisfied by $DivK$ and $curlK$. They return to fall-off estimates for K in their Chapter 11. □

5.3 *Global existence*

The CK proof of global existence for small data is through estimates for the Weyl tensor in a given asymptotically Minkowskian spacetime, estimates of connection and metric of a spacetime with given Weyl tensor, and a bootstrap argument.

5.3.1 *Weyl tensor and Bel Robinson energy*

The fall-off at time infinity of a field in Minkowski spacetime which falls off at space infinity is linked with the conformal invariance of the Minkowski spacetime, and an eventual conformal invariance of the field equations. The conformal invariance is used directly in the conformal methods, it is used indirectly in energy estimates by using the vector field generator of conformal transformations. We have seen that the Einstein equations do not have good properties under conformal transformations, while the Weyl tensor enjoys such properties. We have written in Chapter 8 what we called Bianchi equations for the Riemann tensor, analogous to the Maxwell equations satisfied by the electromagnetic 2-form. The Weyl tensor possesses the same symmetries and antisymmetries as the Riemann tensor and in addition, is traceless. The two tensors coincide when the Ricci tensor vanishes.

We have defined a 4-tensor, the Bel tensor $Q(A)$, associated to a double 2-form A which has the symmetries and antisymmetries of the Riemann tensor

$$Q(A)^{\alpha\beta\lambda\mu} \equiv \frac{1}{2}\{A^{\alpha\rho,\lambda\sigma}A^{\beta..,\mu}_{..\rho,...\sigma} + ({}^*A)^{\alpha\rho,\lambda\sigma}({}^*A)^{\beta..,\mu}_{..\rho,...\sigma}\}.$$

The 4-tensor $Q(A)$ is completely symmetric. It is the analogue for a double 2-form of the Maxwell tensor of the 2-form of electromagnetism.

The Bel–Robinson (BR) energy density associated with vector fields X_1, X_2, X_3, X_4 is the scalar $Q(A)(X_1, X_2, X_3, X_4)$. It is positive if the vectors $X_1, \ldots, X_4$ are causal.

If the tensor A satisfies the Bianchi identities, and is traceless (i.e. $A_{\alpha\beta} := A_\alpha{}^\lambda{}_{,\beta\lambda} = 0$), then $Q(A)$ satisfies the conservation law

$$\nabla_\alpha Q(A)^{\alpha\beta\lambda\mu} \equiv 0. \tag{5.7}$$

A BR energy conservation can be deduced from (5.7) only if the vector fields X_1, X_2, X_3, X_4 are Killing vector fields of the spacetime metric.

If the tensor A satisfies only inhomogeneous Bianchi equations, then $Q(A)$ satisfies an inhomogeneous conservation law (see for example in Chapter 8 the case where A is a non-traceless Riemann tensor)

$$\nabla_\alpha Q(A)^{\alpha\beta\lambda\mu} \equiv L^{\beta\lambda\mu}(A, J), \tag{5.8}$$

with L bilinear in A and the right-hand side J of the Bianchi equations.

In Minkowski spacetime a Killing vector field used to prove global existence of solutions of non-linear wave equations[24] are the scaling and the conformal Killing fields, S and C, which capture space and time fall-off. These vector fields, invariant by translation, are at the origin and in inertial coordinates

$$S_{Mink} := x^i\partial_i + t\partial_t = 0$$
$$C_{Mink} := -(t^2 + r^2)\partial_t - 2x^i\partial_i.$$

CK use vector fields S and C which are approximations of these vector fields to define a BR energy which is not conserved, even for the Weyl tensor, but whose error terms could be computed. However, CK observe that the energy inequality deduced from the identity (5.8) is not sufficient to obtain the time decay required for their global existence proof without the too-strong hypothesis on space fall-off which was an obstacle in the use of conformal techniques. The lack of decay results as usual from the non-vanishing of the ADM mass of the metric. Inspired by the time invariance of the ADM mass, and by its rotational invariance under the asymptotic behaviour assumed of the metric at space infinity, CK replaces estimates of the Weyl tensor by estimates of the Lie derivative, modified to be traceless, with respect to the vector field $T := N^{-1}\partial_t$ and an asymptotic

[24] Morawetz, C. (1961) *Commun. Pres. Appl. Math.*, **14**, 561–8; Klainerman, S. (1986) *Lect. Appl. Math.*, **23**, 293–326.

rotation vector field O, approximation of the vector field dual of the 2-tensor in 3 space

$$O^h_{Mink} := \varepsilon^{ijh}(x^i\partial_j - x^j\partial_i).$$

The traceless tensors $A := \hat{\mathcal{L}}_X W$, $\hat{\mathcal{L}}_X\hat{\mathcal{L}}_Y W$, with $X, Y = T$ or O possess the symmetries and antisymmetries of the Weyl tensor and commute with the Hodge dual star operator (long computations done in Section 7.1 by CK). They satisfy Bianchi equations, but inhomogeneous ones.

CK apply the identities (5.8) to the Bel–Robinson tensors of these modified Lie derivatives of the Weyl tensor, using the approximate conformal Killing vectors of the spacetime they are constructing. They estimate what they call "error terms". However, due to the difference in fall-off in radial and non-radial directions, CK first decompose the tensors into their components in a privileged null frame linked with the geometric, approximately Minkowskian, structure.

5.3.2 *Optical function and null structure*

The CK definition of approximate scale S and conformal C requires a definition of topological 2-spheres replacing the Minkowskian 2-spheres translated from $r =$ constant which are intersections of light cones with the hyperplanes (maximal hypersurfaces in Minkowski spacetime) $t =$ constant. CK introduce what they call an **optical function** u, a solution of the eikonal equation

$$g^{\alpha\beta}\partial_\alpha u\partial_\beta u = 0.$$

The level sets of u are null surfaces Σ_u which intersect the maximal slices M_t of their spacetime along topological 2-spheres $S_{u,t}$ which replace the Euclidean 2-spheres in Minkowski spacetime translated from the spheres $r =$ constant, $t =$ constant. The construction of the optical function is done in three steps, first defined in a neighbourhood of the initial slice by null cones with vertex in the past of the initial slice, then after the last local slice as null surfaces originating from the 2-spheres Σ, and finally by matching the two. The construction and study of the optical function and its higher derivatives is a fundamental point of the CK book.

Assuming the existence of an optical function, CK define a **null structure** of spacetime as a field of null frames composed at each point of two null vectors e_+, e_- orthogonal to $S_{t,u}$, normalized by $(e_+, e_-) = -2$, and two orthonormal vectors tangent to $S_{t,u}$. CK decompose the Weyl tensor W into tensor, vector, and scalar on $S_{t,u}$, analogous with the decomposition of a spacetime p-tensor into t-dependent space tensors of order $\leq p$ of Chapter 6. They define two tensors α, $\underline{\alpha}$ on $S_{t,u}$ by[25]

$$\alpha(X,Y) := W(X, e_+, Y, e_+), \quad \underline{\alpha} := W(X, e_-, Y, e_-),$$

the vectors $\beta, \underline{\beta}$ by

$$\beta(X) := W(X, e_+, e_-, e_+), \quad \underline{\beta} := W(X, e_-, e_+, e_-),$$

[25] The Newman–Penrose formalism also uses such a decomposition, through 2-component spinors.

and the scalar ρ and σ by

$$\rho := W(e_-, e_+, e_-, e_+), \quad \sigma :=^* W(X, e_-, Y, e_-).$$

5.3.3 *Global estimates of the Weyl tensor*

CK constructs through the optical function the approximate conformal Killing fields. Then with long, detailed, and involved computations they provide a "comparison theorem" which estimates various weighted norms for the different components of the Weyl tensors and their derivatives in terms of the Bel–Robinson energies Q_1 and Q_2 associated with causal approximate conformal Killing vector fields of tensors $\hat{\mathcal{L}}_X W$ and $\hat{\mathcal{L}}_X \hat{\mathcal{L}}_Y W$ with $(X, Y) = (O, T)$ They estimate ("boundedness theorem") the "error terms" coming in the Bel–Robinson energy inequalities applied to Q_1 and Q_2 from the right-hand sides appearing in the identities (5.8) when A is one of the considered Lie derivatives. They study separately different components. They obtain estimates, assuming some properties of the deformation tensors associated to the considered Killing vector fields, after many pages of careful computation.

5.3.4 *Connection and metric*

Estimates for the Weyl tensor assume estimates for the second fundamental form and the Ricci tensor of space. Conversely, the connection and Ricci tensor of space can be estimated in a vacuum Einsteinian spacetime in terms of the Weyl tensor by using identities given in Chapters 6 and 8. However, the fall-off estimates require again the distinction between radial and non-radial directions, and the computation takes many pages before CK obtain the results they are looking for.

5.3.5 *Conclusion*

CK deduce from their estimates and a bootstrap argument their global existence theorem for small initial data. They also obtain asymptotic properties for the null components of the curvature tensor of the spacetime they have constructed along a level set $\mathcal{C}_u^+$ of their optical function. Denoting by $\underline{A}$, $\underline{B}$, P and Q, respectively a covariant symmetric 2-tensor, a 1-form and a scalar function on S^2, all depending on u, this asymptotic behaviour is for the components $\underline{\alpha}$, $\underline{\beta}$, ρ, and σ:

$$\lim_{\mathcal{C}_u^+, r\to\infty} r\underline{\alpha} = \underline{A}, \quad \lim_{\mathcal{C}_u^+, r\to\infty} r^2\underline{\beta} = \underline{B}$$

$$\lim_{\mathcal{C}_u^+, r\to\infty} r^3\rho = P, \quad \lim_{\mathcal{C}_u^+, r\to\infty} r^3\sigma = Q.$$

Christodoulou[26] points out that the behaviour of α and β that they find, namely

$$\alpha, \beta = o(r^{-\frac{7}{2}})$$

is inconsistent with a smooth compactification of null infinity. He develops then a rigorous mathematical argument showing that a strengthening of the fall

[26] Christodoulou, D. (2000) *Marcel Grossmann Meeting on General Relativity*, MGIXMM, Gurzadyan (eds. Jantzen and Ruffini), pp. 44–54, World Scientific.

off hypotheses of the initial data which would insure this smooth compactification is incompatible with physical situations and estimates deduced from the quadrupole formula used in weak field approximations [27]. However, both Christodoulou and Damour remark that the regularity of a possible compactification ($C^{1,\frac{1}{2}}$ by the CK theorem) is sufficient for application to actual physical observations.

6 The Klainerman–Nicolo theorem

The work of Klainerman and Nicolo[28], KN, treats the same subject as CK, global existence of spacetimes with small asymptotically Euclidean data, and their asymptotic behaviour. The main tool is the same: Weyl tensor estimates by Bel–Robinson energies, associated with approximate conformal Killing vectors, of some traceless Lie derivatives of the Weyl tensor. A fundamental difference, in addition to technical simplifications, is the replacement by KN of the space-null foliation t, u of CK by a **double null foliation** by outgoing and ingoing null level sets of two optical functions u and $\underline{u}$. KN are thus able to obtain a theorem which gives asymptotic behaviour at null infinity of solutions of the vacuum Einstein equations with asymptotically Euclidean initial data which are required to be small only outside a compact set.

Before stating a theorem, we give a definition. We use the definition (5.1) of CK, but we fix the point x_0, and we denote

$$J_0(M, \bar{g}, K) := Q(x_0; \bar{g}, K)$$

Definition 6.1 *Given a strongly asymptotically Euclidean initial data set $(M, \bar{g}, K)$ and a compact set $\mathbf{C}$, with $M - \mathbf{C}$ diffeomorphic to the exterior of the closed ball of R^3. $J_{\mathbf{C}}$ is defined as follows.*

1. *We define by $\boldsymbol{G}$ the set of all smooth extensions $(\tilde{g}, \tilde{K})$ to the whole of M of the restriction to $\mathbf{C}$ of the data $(\bar{g}, K)$, with $\tilde{g}$ a Riemannian metric and $\tilde{K}$ a symmetric 2-tensor.*
2. *We denote*

$$J_{\mathbf{C}}(M, \bar{g}, K) := \underset{\mathbf{G}}{Inf} J_0(M, \tilde{g}, \tilde{K}).$$

We can now state the first part of the KN theorem.

Theorem 6.2 *(main KN, first part). Consider a strongly asymptotically Euclidean initial data set $(M, \bar{g}, K)$ satisfying the vacuum constraints. There exist a sufficiently large compact set $\mathbf{C}$ and $\varepsilon > 0$ such that if the initial data set $(M, \bar{g}, K)$ satisfies the* ***exterior global smallness condition***

$$J_{\mathbf{C}}(M, \bar{g}, K) < \varepsilon$$

[27] Damour, T. (1986) *Fourth Marcel Grossmann Meeting on General Relativity*, (ed. Ruffini), pp. 365–92, Elsevier.

[28] Klainerman, S. and Nicolo, F. (2002) *The Evolution Problem in General Relativity*, Birkhäuser.

then its maximal vacuum Einsteinian development (V, g) is such that the future V^+ of $M - \mathbf{C}$ can be foliated by a canonical double null foliation whose outgoing leaves are future complete. Similar properties hold for the past V^- of $M - \mathbf{C}$.

KN make the very likely conjecture that their solution is geodesically complete if the full smallness condition is satisfied.

A second part of the KN theorem asserts bounds of norms of various geometric quantities relevant to the problem. We give here only the results concerning the Riemann curvature tensor, the most important physical quantity, as a corollary to the main theorem. They are essentially the same as in the case of the CK theorem, which assumed global smallness of the initial data.

Corollary 6.3 *The null components of the Riemann tensor (compare the conclusion of Section 5.3) admit the following asymptotic behaviour*

$$\underset{V^+}{Supr}^{\frac{7}{2}}|\alpha| \leq C_0, \quad \underset{V^+}{Supr}|\underline{u}|^{\frac{5}{2}}|\underline{\alpha}| \leq C_0,$$

the vectors $\beta, \underline{\beta}$ by

$$\underset{V^+}{Supr}^{\frac{7}{2}}|\beta| \leq C_0, \quad \underset{V^+}{Supr}^{2}|\underline{\beta}| \leq C_0$$

and the scalar ρ and σ by

$$\underset{V^+}{Supr}^{3}|\rho| \leq C_0, \quad \underset{V^+}{Supr}^{3}|\underline{u}|^{\frac{1}{2}}|\sigma| \leq C_0$$

with C_0 numbers depending only on the initial data.

Similar conclusions hold on V^-.

7 The Linblad–Rodnianski theorem

Lindblad and Rodnianski[29] succeeded in proving a result which remained doubtful for a long time, the global existence on R^4 of a solution of the vacuum Einstein equations in wave coordinates for small – i.e. near-Minkowskian – initial data. The Einstein equations in wave coordinates are just a system of quasilinear wave equations, but if one looks simply at these equations, forgetting the harmonicity conditions, they do not satisfy the null condition; it appears that the Minkowski solution should be unstable[30]. Indeed, if one uses the simple iteration scheme

$$g_{\alpha\beta} := \eta_{\alpha\beta} + \varepsilon g^{(1)}_{\alpha\beta} + \varepsilon^2 g^{(2)}_{\alpha\beta} + \dots \tag{7.1}$$

one finds that the vanishing of the terms in ε of the expansion of the Einstein equations in wave coordinates reduces to the wave equation in the Minkowski metric for $g^{(1)}_{\alpha\beta}$,

$$\Box_\eta g^{(1)}_{\alpha\beta} = 0, \tag{7.2}$$

[29] Lindblad, H. and Rodnianski, I. (2005) *Commun. Math. Phys.*, **246**(1), 43–110 and arXiv:math/0411109v1.

[30] Choquet-Bruhat, Y. (1973) *C. R. Acad. Sci. Paris*, **276**A, 281–4.

the terms in ε^2 are a system of wave equations with sources quadratic in the derivatives of the $g^{(1)}_{\alpha\beta}$,

$$\Box_\eta g^{(2)}_{\alpha\beta} = q_{\alpha\beta}(\partial g^{(1)}, \partial g^{(1)}). \tag{7.3}$$

For some initial data decaying in space like r^{-1} there are solutions of (7.2) which decay in time only[31] like t^{-1}. Integrating (7.3) implies that, for such initial data, $g^{(2)}_{\alpha\beta}$ decays only like $t^{-1}\log t$, in disagreement with the then expected behaviour in the case of the stability of Minkowski spacetime. It was suggested, though never proved, that stability holds, but only for initial data satisfying the constraints.

Lindblad and Rodnianski (LR) take another route to prove the stability theorem: they couple the Einstein equations in wave coordinates with the wave coordinates conditions. We saw in Chapter 11 that the Einstein equations do not quite satisfy the polarized null condition, and this fact signals difficulties in proving global existence for near-Minkowskian data, even when taking into account the harmonicity conditions. We state the Lindblad–Rodnianski global existence theorem and indicate briefly some idea of their proof.

7.1 *The Einstein equations in wave coordinates*

We work on R^4; the coordinates are global, ∂_α denotes the partial derivative $\partial/\partial x^\alpha$.

The harmonicity conditions $g^{\alpha\beta}\Gamma^\lambda_{\alpha\beta} = 0$ can equivalently be written

$$g^{\alpha\beta}\partial_\beta g_{\alpha\mu} = \frac{1}{2} g^{\alpha\beta}\partial_\mu g_{\alpha\beta}. \tag{7.4}$$

We have seen that in wave coordinates, for a scalar function, the covariant wave operator reads

$$\Box_g \equiv g^{\alpha\beta}\partial^2_{\alpha\beta} := \tilde{\Box}_g. \tag{7.5}$$

Setting

$$g_{\mu\nu} := \eta_{\mu\nu} + h_{\mu\nu},$$

the Einstein equations in wave coordinates are of the form

$$\tilde{\Box}_g h_{\mu\nu} + Q_{\mu\nu}(g)(\partial h, \partial h) = 0. \tag{7.6}$$

The following lemma is proved by a straightforward but long computation[32].

Lemma 7.1 *The quadratic form Q in ∂h has the following structure (indices raised with the Minkowski metric η)*

$$Q_{\mu\nu}(g)(\partial h, \partial h) \equiv P_{\mu\nu}(\partial h, \partial h) + A_{\mu\nu}(\partial h, \partial h) + B(h)(\partial h, \partial h)$$

[31] Such initial data have a positive time derivative; the asymptotic behaviour is computed through the Kirchhoff formula.

[32] Compare Section XI.8.

with

$$P_{\mu\nu}(\partial h, \partial h) \equiv \frac{1}{2}\left(\partial_\mu h_{\alpha\beta}\partial_\nu h^{\alpha\beta} - \frac{1}{2}\partial_\mu h^\alpha_\alpha \partial_\nu h^\beta_\beta\right)$$

while for each pair of indices μ, ν, the quadratic form $A_{\mu\nu}$ is a linear combination of the standard null forms in ∂h. The coefficients of the quadratic form B vanish for $h = 0$, B is a cubic term.

We see that only the quadratic form $P_{\mu\nu}$ does not satisfy the null condition. However, $tr_\eta h \equiv h^\alpha_\alpha$ satisfies the equation

$$\tilde{\Box}_g tr_\eta h + \eta^{\mu\nu} Q_{\mu\nu}(g)(\partial h, \partial h),$$

whose right-hand side obeys the null condition.

7.2 *Initial data*

LR consider an initial data set $(R^3, \bar{g}, K)$ satisfying the vacuum Einstein constraints and which is of the form[33]

$$\bar{g} = \delta + \bar{h}^0 + \bar{h}^1 \quad \text{with} \quad \bar{h}^0_{ij} \equiv \chi(r)\frac{m}{r}\delta_{ij} \quad \text{and}$$
$$\bar{h}^1 = o(r^{-1-\alpha}), \quad K = o(r^{-2-\alpha}), \ \alpha > 0,$$

where χ is a given smooth function equal to 1 for large r and equal to 0 for small r. The number m is the ADM mass of the initial data $\bar{g}$.

LR construct analysis initial data, $g_{\alpha\beta,0}$ and $(\partial_t g_{\alpha\beta})_{,0}$ satisfying the initial harmonicity conditions as we did in the case of the local Cauchy problem, taking $g_{ij,0} := \bar{g}_{ij}$ and $g_{i0,0} := 0$ but they find it convenient to take for the initial lapse the function

$$N_0^2 := 1 - \chi(r)\frac{m}{r}$$

The initial values $(\partial_t g_{ij})_{,0}$ are determined by the tensor K,

$$(\partial_t g_{ij})_{,0} := -2N_0^2 K_{ij};$$

$(\partial_t g_{\alpha\beta})_{,0}$ and $(\partial_t g_{\alpha\beta})_{,0}$ are determined by the harmonicity conditions as in Chapter 6.

7.3 *Unknowns and norms*

In order to get rid of the problem caused by the ADM mass m, LR take as new unknown the tensor h^1 defined by

$$h^1_{\alpha\beta} := h_{\alpha\beta} - h^0_{\alpha\beta}, \qquad h_{\alpha\beta} := g_{\alpha\beta} - \eta_{\alpha\beta}$$

[33] In a previous paper LR took initial data which are exactly Schwarzschild outside of a compact set.

with

$$h^0_{\alpha\beta}(t,.) := \chi\left(\frac{r}{t}\right)\chi(r)\frac{m}{r}\delta_{\alpha\beta}$$

where $s \mapsto \chi(s)$ is a smooth function of a real variable equal to 1 for $s \geq \frac{3}{4}$, equal to 0 for $s \leq \frac{1}{2}$. Then

$$\chi\left(\frac{r}{t}\right) = 1, \quad r \geq \frac{3t}{4} \quad \text{and} \quad \chi\left(\frac{r}{t}\right) = 0, \quad r \leq \frac{t}{2}.$$

Expressed in terms of h^1 the Einstein equations in wave coordinates are

$$\tilde{\Box}_g h^1_{\mu\nu} + \tilde{\Box}_g h^0_{\mu\nu} + Q_{\mu\nu}(g)(\partial h, \partial h) = 0. \tag{7.7}$$

The advantage of the splitting of h is that h^1 contains no mass term, i.e. falls off more rapidly than r^{-1}, while the flat three-dimensional Laplacian of $\frac{1}{r}$ being identically zero when $r \neq 0$, the term $\tilde{\Box}_g h^0_{\mu\nu}$ is manageable.

LR define, like CK, weighted norms for tensors on R^4, but they use the Minkowski metric itself and its conformal Killing vector fields – they don't need to introduce approximate such vectors, nor an optical function which was a hard part of the CK proof. LR denote by $\mathcal{Z}$ the following set of conformal Killing vector fields of the Minkowski spacetime M^4 (translations, rotations, and scale)

$$\mathcal{Z} := \{Z_{(1)}, \ldots Z_{(11)}\} \equiv \{\partial_\alpha,\ x^\alpha\partial_\beta - x^\beta\partial_\alpha,\ x^\alpha\partial_\alpha\}.$$

The norm at time t of the spacetime tensor h^1 is

$$\mathcal{E}_n(t) := \sum_{|I|\leq n, Z\in\mathcal{Z}} ||w^{\frac{1}{2}|}\partial Z^I h^1(t,.)||_{L^2}$$

where, with $\gamma, \mu > 0$

$$w := 1 + (1 + |r-t|)^{1+2\gamma} \quad \text{if } r - t > 0$$
$$w := 1 + (1 + |r-t|)^{-2\mu} \quad \text{if } r - t < 0$$

where $I = I_1 \ldots I_{11}$ is a multi-index, $Z^I := Z^{I_1}_{(1)} \ldots Z^{I_{11}}_{(11)}$, $|I| = I_1 + \ldots + I_{11}$.

The smallness of initial data is characterized by the smallness of the sum $\mathcal{E}_n(0) + m$.

Remark 7.2 *$\mathcal{E}_n(0)$ is equivalent to, with $\bar{\partial}$ denoting space* derivatives,

$$E_n(0) := \sum_{0\leq|I|\leq n} ||(1+r)^{\frac{1}{2}+\gamma+|I|}\bar{\partial}^I\bar{h}^1||_{L^2} + ||(1+r)^{\frac{1}{2}+\gamma+|I|}\bar{\partial}^I K||_{L^2}.$$

7.4 *LR theorem*

We have seen that if the initial data for $g_{\alpha\beta}$ satisfy the Einstein constraints and the wave coordinate conditions then there exists a local in time solution of the

reduced Einstein equations, written here $\eta_{\alpha\beta} + h_{\alpha\beta}$, which satisfies the wave coordinate conditions, hence the full Einstein equations.

LR state their global existence theorem as follows.

Theorem 7.3 *There exists a constant $\varepsilon_0 > 0$ such that if $\varepsilon \leq \varepsilon_0$ and the initial data $h^1|_{t=0}, \partial_t h^1|_{t=0}$ are smooth and obey $\mathcal{E}_n(0) + m \leq \varepsilon$ together with the condition*

$$\lim_{|x|\to\infty} \inf |h^1(0,x)| \to 0,$$

then the solution of the reduced Einstein equations can be extended to a global smooth solution, geodesically complete, with h^1 such that

$$\mathcal{E}_n(t) \leq C_n \varepsilon (1+t)^{c\varepsilon},$$

where C_n is a constant depending only on n and c is independent of ε.

The proof is by proving the non-blow-up of the norm $\mathcal{E}_n(t)$ for any finite t. This proof rests on a bootstrap argument, after many delicate estimates using Klainerman weighted inequalities which generalize the usual Sobolev inequalities, and what they call Hardy-type inequalities, with weight depending on the distance to the Minkowskian light cone[34]

In their paper Lindblad and Rodnianski extend their global existence result to the vacuum Einstein-scalar case. Loizelet[35] proves a similar theorem for the electrovac case, using wave coordinates for the metric and Lorentz gauge for the electromagnetic potential.

[34] In the classical Hardy inequality on R^3 the weight depends on the distance to the origin. Similar inequalities on asymptotically Euclidean manifolds are stated in Appendix I.

[35] Loizelet, J. (2005) *C. R. Acad. Sci. Paris*, Serie I, **340**.

XVI

GLOBAL EXISTENCE THEOREMS THE COSMOLOGICAL CASE

1 Introduction

It has become usual to call "cosmological" those Einsteinian spacetimes which have compact spacelike sections, in contradistinction to those spacetimes which have asymptotically Euclidean space sections and are used to model the motions of isolated bodies. However, the cosmos we live in may well have non-compact spacelike sections. In fact very little is known about our universe as a whole. It is legitimate for the mathematician to study all possible models, with arbitrary topology as well as arbitrary Lorentzian metric. It is still believed by the majority of physicists and astronomers that such a manifold does model the cosmos, and moreover its metric satisfies the Einstein equations, eventually in dimensions greater than four. This hypothesis opens a vast field of investigations, where remarkable conjectures have been proposed. Many results have been obtained, sometimes surprising, but many fundamental questions remain open.

Among the most theoretically interesting questions debated today are the questions of the beginning and the end of our Universe. It is generally believed that it had a beginning. Most modern astronomers tell us the beginning was a Big Bang, that is, for the mathematician, a singularity of the spacetime. The ultimate future fate of the cosmos we live in cannot now really be predicted, in spite of the most recent observational data. For the mathematician the question is of the future global existence of a solution of the Einstein equations. Since the cosmos is a physical as well as a geometrical object, where the elapsed time depends on the trajectory of the observer, the definition of the word "global" in the context of Einstein equations requires some thought. Since observers in free fall follow geodesics, the requirement for the future global existence of an Einsteinian spacetime is usually taken as equivalent to the completeness of all future-directed causal geodesics.

It is conjectured in the $3+1$ case that the only vacuum Einsteinian spacetimes with compact spacelike sections which are complete both towards the past and towards the future are the product of R by a locally flat manifold. Global existence in one time direction, let us say the future, has been proved over the years for solutions with a decreasing number of symmetries. Global future existence for solutions of the Cauchy problem for the vacuum Einstein equations with large initial data is known only in cases where the isometry group is two-dimensional

(see the next section by V. Moncrief). The other global existence proofs require small initial data. It is still thought possible to find a proof of global existence for classes of S^1 symmetric large initial data, while the singularity theorems make it unlikely for generic data.

2 Gowdy cosmological models

Section contributed by Vincent Moncrief

Stationary axisymmetric vacuum metrics are an important class of solutions to Einstein's field equations which, in the asymptotically flat case, can be used to model the exteriors of rotating stars. They are characterized as having two commuting Killing fields, one spacelike and the other timelike. In 1971, R. Gowdy introduced a cosmological analogue to the stationary axisymmetric problem by initiating the study of vacuum metrics on spatially compact manifolds that admit two commuting, spacelike Killing fields[1]. He showed that interesting families of such spacetimes are definable on the manifolds $T^3 \times R$ $S^3 \times R$ and $S^2 \times S^1 \times R$ and presented explicit solutions of the field equations in some special cases. Gowdy metrics provide perhaps the simplest examples of spatially inhomogeneous solutions describing gravitational radiation in a closed universe and they have often been used to test ideas about cosmic censorship and the nature of cosmological singularities[2,3].

For simplicity let us restrict our attention to Gowdy metrics defined on $T^3 \times R$ taking periodic coordinates $\{x^a\} = \{x^1, x^2\}$ and $x^3 = \theta$ on T^3, writing $x^0 = t$ for the time coordinate and setting

$$ {}^{(4)}g_{\mu\nu}\, dx^\mu\, dx^\nu = \exp[2a(t,\theta)](-dt^2 + d\theta^2) + g_{ab}(t,\theta)\, dx^a\, dx^b \qquad (2.1) $$

so that $\partial/\partial x^1$ and $\partial/\partial x^2$ are spacelike (commuting) Killing fields. The Riemannian 2-metric g_{ab} determines the intrinsic geometry of the integral surfaces of the Killing fields $\{\partial/\partial x^1, \partial/\partial x^2\}$ and, upon adopting for convenience Gowdy's parameterization, can be expressed as

$$ (g_{ab}) = R \begin{bmatrix} \cosh W + \cos\Phi \sinh W & \sin\Phi \sinh W \\ \sin\Phi \sinh W & \cosh W - \cos\Phi \sinh W \end{bmatrix} \qquad (2.2) $$

in terms of functions $\{R, W, \Phi\}(t,\theta)$ of t and θ alone. Note that the area element of the 2-metric g_{ab} is then given by

$$ \{\det g_{ab}\}^{1/2} = R(t,\theta)\,. \qquad (2.3) $$

From the vacuum field equations one gets

$$ g^{ab}\, {}^4R_{ab} = \frac{e^{-2a}}{R}\left(\frac{\partial^2 R}{\partial t^2} - \frac{\partial^2 R}{\partial \theta^2}\right) = 0 \qquad (2.4) $$

[1] Gowdy, R. (1971) *Phys. Rev. Lett.*, **27**, 826–9 and erratum p. 1102. See also by the same author (1974) *Ann. Phys.*, **83**, 203–41.

[2] Isenberg, J. and Moncrief, V. (1990) *Ann. Phys.*, **199**, 84–122.

[3] Chruściel, P., Isenberg, J., and Moncrief, V. (1990) *Class. Quant. Grav.*, **7**, 1671–80.

and thus finds that R must satisfy the wave equation on $S^1 \times R$. However, Gowdy showed, by considering transformations of the $\{t, \theta\}$ coordinates on this cylinder which preserve the metric form, that he could justify choosing the particular solution $R(t, \theta) = t$ without any essential loss of generality. This corresponds to choosing a time function that is constant on the 2-surfaces generated by the Killing fields and takes a value on each such surface that is proportional to the area of that surface.

The remaining vacuum field equations correspond to the non-linear evolution equations for $\{W, \Phi\}(t, \theta)$ derivable as Euler–Lagrange equations from the Lagrangian

$$L = \frac{1}{2} \int_{S^1} d\theta \left\{ t \left[W_{,t}^2 - W_{,\theta}^2 + \sinh^2 W (\Phi_{,t}^2 - \Phi_{,\theta}^2) \right] \right\} = \int_{S^1} d\theta \, \mathbf{L} \tag{2.5}$$

namely,

$$W_{,tt} + \frac{1}{t} W_{,t} - W_{,\theta\theta} - \sinh W \cosh W (\Phi_{,t}^2 - \Phi_{,\theta}^2) = 0 \,, \tag{2.6}$$

$$\sinh W \left(\Phi_{,tt} + \frac{1}{t} \Phi_{,t} - \Phi_{,\theta\theta} \right) + 2 \cosh W (\Phi_{,t} W_{,t} - \Phi_{,\theta} W_{,\theta}) = 0 \tag{2.7}$$

and to constraint equations for the remaining metric function $a(t, \theta)$ expressible as

$$2 \tilde{a}_{,\theta} = t (W_{,t} W_{,\theta} + \sinh^2 W \, \Phi_{,t} \, \Phi_{,\theta}) =: \mathbf{P} \tag{2.8}$$

$$2 \tilde{a}_{,t} = \frac{t}{2} (W_{,t}^2 + W_{,\theta}^2 + \sinh^2 W (\Phi_{,t}^2 + \Phi_{,\theta}^2)) =: \mathbf{H} \tag{2.9}$$

$$2(\tilde{a}_{,tt} - \tilde{a}_{,\theta\theta}) = -\frac{1}{2} \left[W_{,t}^2 - W_{,\theta}^2 + \sinh^2 W (\Phi_{,t}^2 - \Phi_{,\theta}^2) \right] =: -\frac{\mathbf{L}}{t} \tag{2.10}$$

where

$$\tilde{a} = a + 1/4 \, \ln(t) \,. \tag{2.11}$$

The constraint (2.8) forces, upon integration over the circle, the vanishing of the total momentum

$$\int_{S^1} d\theta \, [t (W_{,t} W_{,\theta} + \sinh^2 W \, \Phi_{,t} \, \Phi_{,\theta}] = \int_{S^1} d\theta \mathbf{P} = 0 \tag{2.12}$$

but this additional integrability condition is easily shown to be conserved by the evolution equations, which in turn imply the redundancy of (2.10), which follows from the remaining equations. One can regard $\mathbf{L}$, $\mathbf{H}$, and $\mathbf{P}$ defined in the formulae above as Lagrangian, Hamiltonian, and momentum densities respectively for the gravitational waves.

Noting that $\{d W^2 + \sinh^2 W \, d \Phi^2\}$ is a particular representation for the standard metric on hyperbolic 2-space H^2 one sees that the action for the evolution equations is nearly that for a wave map from $S^1 \times R$ to H^2 and would be exactly

so except for the factor of t in the expression for L. The presence of this factor, which vanishes in the limit that the Killing 2-surfaces tend to zero area, allows the solutions, which otherwise would be globally regular[4], to develop singularities as $t \to 0$. Note that in the special case for which $\Phi =$ constant (the so-called "polarized" case) the evolution equation for W reduces to a linear wave equation which can be solved explicitly. Its solutions are (semi-) globally defined on the time interval $(0, \infty)$ and generically blow up at the boundary $t \to 0$, which then corresponds to a (curvature singular) Big Bang[2].

For the non-linear problem a corresponding local and (semi-) global existence theorem was established by Moncrief[5] which showed that sufficiently smooth initial data prescribed at $t = t_0 > 0$ (and satisfying the integrability condition (2.12) always determines a unique solution of the field equations which remains non-singular on the entire time interval $(0, \infty)$ but which develops a "crushing singularity" of divergent mean curvature at the boundary corresponding to $t \to 0$. It follows from a theorem of Hawking that this crushing singularity lies at the boundary of the maximal Cauchy development of the given initial data. It was also shown in[5] that the (timelike) curves orthogonal to the level surfaces of Gowdy's time function t always tend to infinite length in the limit as $t \to \infty$.

While the arguments given in[5] are somewhat technical the essential estimate can be simplified by the following observation[6]. Using Duhamel's principle to solve the inhomogeneous wave equation (2.10) for $\tilde{a}(t, \theta)$ and differentiating the resulting integral expression for $\tilde{a}(t, \theta)$ with respect to t and appealing to (2.9) one derives a formula for the Hamiltonian density $\mathbf{H}$, evaluated at an arbitrary point p within the domain of local existence, in terms of an integral of the Lagrangian density $\mathbf{L}$ over the (one-dimensional) light cone from p to the initial surface at $t = t_0$ supplemented by an explicit contribution (a solution of the homogeneous wave equation) computable from the initial data. With the resulting integral equation one can easily apply Gronwall's lemma to prove that $\mathbf{H}(t, \theta)$ cannot blow up pointwise on the interval $(0, \infty)$ and thus that the first derivatives of the evolving fields $\{W, \Phi\}(t, \theta)$ cannot blow up on this interval. Armed with such bounds it is straightforward to show that Sobolev norms (of any desired order as determined by the smoothness of the given data) cannot blow up on this interval and hence (by direct integration) that $\tilde{a}(t, \theta)$ also remains well-behaved for all $t > 0$.

While these early, rough arguments establish (semi-) global existence for Gowdy spacetimes, they leave unanswered many additional questions about the asymptotic behaviours of the solutions in both the expanding ($t \to \infty$) and the collapsing ($t \to 0$) directions. More recently, however, H. Ringström has succeeded in giving a detailed characterization of this limiting behaviour (in both temporal directions) and, in particular, in proving that (strong) cosmic

[4] Gu Chao Hao (1980) *Commun. Pures Appl. Math.*, **33**, 727–37.
[5] Moncrief, V. (1981) *Ann. Phys.*, **132**, 87–107.
[6] Chruściel, P. (1992) *Contemp. Math.*, **132**, Amer. Math. Soc., Providence, RI.

censorship holds for the class of Gowdy spacetimes on $T^3 \times R$ discussed above[7]. Since the precise statements of Ringström's results are somewhat technical we shall not describe them in detail here. Among many other things, though, his results provide a rigorous foundation for some formal asymptotic expansions given by B. Grubišić and V. Moncrief in[8] and extend the rigorous Fuchsian analysis results of S. Kichenassamy and A. Rendall given in[9] to cover generic solutions whose big bang asymptotics can involve the "spikes" discovered numerically by B. Berger in[10].

Most work on Gowdy spacetimes has made use of the Gowdy time function t defined above, but work by Isenberg and Moncrief has shown that *global* CMC (i.e. constant mean curvature) slicings also exist for these spacetimes[11]. A somewhat larger class of metrics defined on $T^3 \times R$ that still admit two commuting, spacelike Killing fields, but for which Gowdy's choice of time function is no longer available, has been studied by Berger, Isenberg, Chruściel and Moncrief[12], who proved a (semi-) global existence theorem for these spacetimes though the use of light-cone and energy estimates.

3 S^1 invariant Einsteinian universes, equations

3.1 *Introduction*

We treat in some detail the global existence proof for classes of cosmological spacetimes with S^1 spacelike isometry group for its intrinsic interest (see Section 4.6), but also as an example of the various geometrical and analytical methods which appear in the study of global problems for the Einstein equations.

This section and the following one are based on work in collaboration with V. Moncrief refered to below as[13] CBM1, CBM2, and CB3.

3.2 *Definition*

The vacuum Einstein equations for a spacetime with spatial S^1 isometry group are locally formally similar to the vacuum Einstein equations on a stationary spacetime obtained in Chapter 14, if we assume that the isometry group S^1

[7] Ringström, H. (2004) *Commun. Pure Appl. Math.*, **57**(5), 657–703; and *Math. Proc. Cambridge Philos. Soc.*, **136**(2), 485–512; *Class. Quant. Grav.*, **21**(3), S305–S322; (2006) *Commun. Pure Appl. Math.*, **59**, 977–1041.

[8] Grubišić, B. and Moncrief, V. (1993) *Phys. Rev. D*, **47**(6), 2371–82.

[9] Kichenassamy, S. and Rendall, A. (1998) *Class. Quant. Grav.*, **15**(5), 1339–55.

[10] Berger, B. and Moncrief, V. (1993) *Phys. Rev. D*, **48**(10), 4676–87.

[11] Isenberg, J. and Moncrief, V. (1982) *Commun. Math. Phys.*, **86**, 485–93.

[12] Berger, B., Chruściel, P., Isenberg, J., and Moncrief, V. (1997) *Ann. Phys.*, **260**, 117–48.

[13] CBM1: Choquet-Bruhat, Y. and Moncrief, V. (1996) Existence theorem for solutions of Einstein equations with 1 parameter spacelike isometry group. *Proc. Symp. Pure Math.*, **59** (eds. H. Brezis and I. E. Segal), pp. 67–80.

CBM2: Choquet-Bruhat, Y. and Moncrief, V. (2001) Future global in time Einsteinian spacetimes with $U(1)$ isometry group. *Ann. Henri. Poincaré*, **2**, 1007–64.

CB3: Choquet-Bruhat, Y. (2004) in *Einstein Equations, Large Scale Behavior of Gravitational Fields* (eds. H. Friedrich and P. Chruściel), Birkhaüser.

which leaves the spacetime metric $(V, {}^{(4)}g)$ invariant endows the manifold V with the structure of a principal fibre bundle with typical fibre and Lie group S^1. The fibres are the orbits of the group, supposed now to be spacelike; the base is of the form $\Sigma \times R$, with Σ a smooth orientable two-dimensional manifold. The metric ${}^{(4)}g$ can be written on V under the form adapted to the bundle structure

$$ {}^{(4)}g = \tilde{g} + e^{2\gamma}(\theta)^2, $$

where γ and $\tilde{g}$ can be identified respectively with a scalar function and a Lorentzian metric on the base manifold $\Sigma \times R$. The 1-form θ on V is an S^1 connection: it reads in a local trivialization of the bundle $V \to \Sigma \times R$, with x^3 a coordinate on the fibre S^1 (i.e. with $x^3 = 0$ and $x^3 = 2\pi$ identified) and $(x^\alpha) \equiv (x^a, t)$, $a = 1, 2$ coordinates on $\Sigma \times R$, quotient of V by the isometry group.

$$ \theta = dx^3 + A_\alpha dx^\alpha. $$

$A := A_\alpha dx^\alpha$ is only a locally defined 1-form on $\Sigma \times R$, it depends on the trivialization. The curvature of the S^1 connection is represented by a globally defined 2-form

$$ F = dA. $$

The Lorentzian metric $\tilde{g}$ is globally defined on $\Sigma \times R$.

3.3 *Equations*

The Kaluza–Klein identities written in Chapter 14 for spacetimes with 1-parameter isometry group give the following equations[14] for a solution of the vacuum Einstein equations ${}^{(4)}R_{\alpha\beta} = 0$

$$ {}^{(4)}R_{\alpha\beta} \equiv \tilde{R}_{\alpha\beta} - \frac{1}{2}e^{2\gamma}F_\alpha{}^\lambda F_{\beta\lambda} - \tilde{\nabla}_\alpha\partial_\beta\gamma - \partial_\alpha\gamma\partial_\beta\gamma = 0, \tag{3.1} $$

$$ {}^{(4)}R_{a3} \equiv -\frac{1}{2}e^{-\gamma}\tilde{\nabla}_\beta(e^{3\gamma}F_\alpha{}^\beta) = 0, \tag{3.2} $$

$$ {}^{(4)}R_{33} \equiv -e^{2\gamma}\left\{-\frac{1}{4}e^{2\gamma}F_{\alpha\beta}F^{\alpha\beta} + g^{\alpha\beta}\partial_\alpha\gamma\partial_\beta\gamma + g^{\alpha\beta}\tilde{\nabla}_\alpha\partial_\beta\gamma\right\} = 0, \tag{3.3} $$

where a tilde denotes geometric objects defined through the metric $\tilde{g}$.

3.4 *Twist potential*

Equations (3.2) express the vanishing of the codifferential in the metric $\tilde{g}$; of the 2-form $e^{3\gamma}F$; that is, equivalently, the vanishing of the differential of the adjoint 1-form

$$ E := e^{3\gamma *}F, \quad ({}^*F)_\alpha \equiv \frac{1}{2}\tilde{\eta}_{\alpha\beta\lambda}F^{\beta\lambda}. \tag{3.4} $$

[14] CB-DM1. Equations obtained in Hamiltonian form by Moncrief, V. (1986) *Ann. Phys.*, **167**(1), 118–42.

Conversely

$$F_{\alpha\beta} \equiv e^{-3\gamma}\tilde{\eta}_{\alpha\beta\lambda}E^{\lambda}. \tag{3.5}$$

The general solution of the equation

$$dE = 0 \tag{3.6}$$

on $V_3 := \Sigma \times R$ is, if Σ is compact

$$E = d\omega + H, \tag{3.7}$$

where ω is a scalar function on V_3, called the **twist potential**, and H a representative of a 1-cohomology class of V_3. In all that follows we take $H = 0$. Extending the global existence result to $H \neq 0$ is a non-trivial open problem.

The equations ${}^{(4)}R_{a3} = 0$ are equivalent (when H is chosen zero) on $(V_3, \tilde{g})$ to a quasilinear wave equation for ω, depending on γ and its first derivatives.

3.5 *Wave map system*

When F is replaced by its value in terms of ω, the equation ${}^{(4)}R_{33} = 0$ is a quasilinear wave equation for γ, depending on ω and its first derivatives. It turns out that if the metric $\tilde{g}$ is weighted by the conformal factor $e^{-2\gamma}$, i.e. we set ${}^{(3)}g := e^{2\gamma}\tilde{g}$, then the 2-form F is linked to E by $F_{\alpha\beta} \equiv e^{-4\gamma}{}^{(3)}\eta_{\alpha\beta\lambda}E^{\lambda}$, and the equations ${}^{(4)}R_{33} = 0$ and ${}^{(4)}R_{\alpha 3} = 0$ take the form

$${}^{(3)}\nabla^{\alpha}\partial_{\alpha}\gamma + \frac{1}{2}e^{-4\gamma}{}^{(3)}g^{\alpha\beta}\partial_{\alpha}\omega\partial_{\beta}\omega = 0, \tag{3.8}$$

$${}^{(3)}\nabla^{\alpha}\partial_{\alpha}\omega - 4{}^{(3)}g^{\alpha\beta}\partial_{\alpha}\omega\partial_{\beta}\gamma = 0, \tag{3.9}$$

with ${}^{(3)}\nabla$ the covariant derivatives on $\Sigma \times R$ in the metric ${}^{(3)}g$.

Lemma 3.1 *Equations (3.8) and (3.9) for the pair $u \equiv (\gamma, \omega)$ are the equations of a wave map from $(\Sigma \times R, {}^{(3)}g)$ into the Poincaré plane (R^2, G),*

$$G = 2(d\gamma)^2 + (1/2)e^{-4\gamma}(d\omega)^2.$$

Proof Straightforward computation: the non-zero Christoffel symbols of the metric G are

$$G^1_{22} \equiv G^{\gamma}_{\omega\omega} = \frac{1}{2}e^{-4\gamma}, \quad G^2_{12} = G^2_{21} = G^{\omega}_{\gamma\omega} = -2.$$

□

Remark 3.2 *The scalar and Riemann curvatures of the metric G are*

$$R_{12,12} = -2e^{-4\gamma}, \quad R = -4, \tag{3.10}$$

The Poincaré plane is a Riemannian manifold of constant negative curvature.

Remark 3.3 *A similar result for stationary spacetimes is the Ernst equation, a harmonic map, which admits various formulations depending on the chosen coordinates.*

3.6 *Three-dimensional Einstein equations*

One deduces by computations using the formulae linking the Ricci tensors of two conformal metrices (see Appendix VI) that, when F is replaced by its value in terms of ω

$$^{(4)}R_{\alpha\beta} +^{(4)} g_{\alpha\beta}{}^{(4)}R_{33} \equiv^{(3)} R_{\alpha\beta} - \rho_{\alpha\beta}, \quad \text{with} \quad \rho_{\alpha\beta} \equiv \partial_\alpha u.\partial_\beta u,$$

where a dot denotes a scalar product in the metric of the Poincaré plane; that is,

$$\partial_\alpha u.\partial_\beta u := 2\partial_\alpha\gamma\partial_\beta\gamma + \frac{1}{2}e^{-4\gamma}\partial_\alpha\omega\partial_\beta\omega. \tag{3.11}$$

We have obtained the following lemma

Lemma 3.4 *When* $^{(4)}R_{3\alpha} = 0$ *and* $^{(4)}R_{33} = 0$ *the Einstein equations* $^{(4)}R_{\alpha\beta} = 0$ *are equivalent to Einstein equations on the 3-manifold* $V_3 := \Sigma \times R$ *for the metric* $^{(3)}g$ *with source the stress energy tensor of the wave map* $u = (\gamma, \omega)$.

$$^{(3)}R_{\alpha\beta} = \partial_\alpha u.\partial_\beta u. \tag{3.12}$$

We write the Lorentzian metric $^{(3)}g$ in the usual $2+1$ form

$$^{(3)}g = g_{\alpha\beta}\theta^\alpha\theta^\beta \equiv -N^2(\theta^0)^2 + g_{ab}\theta^a\theta^b,$$

with

$$\theta^0 \equiv dt, \quad \theta^a \equiv dx^a + \nu^a dt.$$

The scalar N and the vector ν are respectively the lapse and the shift of $^{(3)}g$, and

$$g = g_{ab}dx^a dx^b$$

is a Riemannian metric on Σ, depending on t.

In three dimensions the Einstein equations are essentially non-dyamical[15], but not quite as we show now. We consider a $2+1$ decomposition of the equations and we look for a solution $(\bar{g}, k, N, \nu)$ on $\Sigma \times R$.

3.6.1 *Constraints on* Σ_t

One denotes by k the extrinsic curvature of Σ_t as a submanifold of $(\Sigma \times R,^{(3)} g)$. Then, with ∇ the covariant derivative in the metric g,

$$k_{ab} \equiv (2N)^{-1}(-\partial_t g_{ab} + \nabla_a\nu_b + \nabla_b\nu_a). \tag{3.13}$$

The constraints, equations to be satisfied on each Σ_t, are the momentum constraint,

$$^{(3)}R_{0a} \equiv N(-\nabla_b k^b_a + \partial_a\tau) = \partial_0 u.\partial_a u, \tag{3.14}$$

[15] See the study of $2+1$ Einstein gravity by Andersson, L., Moncrief, V., and Tromba, A. (1997) *J. Geom. Phys.*, **23**(3,4), 1991–2005.

and the Hamiltonian constraint,

$$2N^{-2(3)}S_{00} \equiv R(g) - k_b^a k_a^b + \tau^2 = N^{-2}\partial_0 u.\partial_0 u + g^{ab}\partial_a u.\partial_b u, \tag{3.15}$$

The constraints do not contain second derivatives transversal to Σ_t of g or u. To transform the constraints into an elliptic system one uses the conformal method (see Chapter 7). We set

$$g_{ab} = e^{2\lambda}\sigma_{ab},$$

where σ is a Riemannian metric on Σ, depending on t, on which we will comment later, and

$$k_{ab} = h_{ab} + \frac{1}{2}g_{ab}\tau$$

where τ is the g-trace of k; hence h is traceless.

We denote by D a covariant derivation in the metric σ. We set

$$u' = N^{-1}\partial_0 u$$

with ∂_0 the Pfaff derivative of u, namely

$$\partial_0 = \frac{\partial}{\partial t} - \nu^a\partial_a \quad \text{with} \quad \partial_a = \frac{\partial}{\partial x^a}$$

and

$$\dot{u} = e^{2\lambda}u'.$$

The momentum constraint on Σ_t reads if τ is constant in space, a choice which we will make

$$D_b h_a^b = L_a, \quad \text{with} \quad L_a \equiv -D_a u.\dot{u} \tag{3.16}$$

This is a linear equation for h, with left-hand side independent of λ. The general solution is the sum of a transverse traceless tensor $h_{TT} \equiv q$ and a conformal Lie derivative r. Such tensors are L^2-orthogonal on (Σ, σ).

The Hamiltonian constraint reads as the semilinear elliptic equation in λ :

$$\Delta\lambda = f(x,\lambda) \equiv p_1 e^{2\lambda} - p_2 e^{-2\lambda} + p_3, \tag{3.17}$$

with $\Delta \equiv \Delta_\sigma$ the Laplacian in the metric σ and:

$$p_1 \equiv \frac{1}{4}\tau^2, \quad p_2 \equiv \frac{1}{2}(|\ \dot{u}\ |^2 + |\ h\ |^2), \quad p_3 \equiv \frac{1}{2}(R(\sigma) - |Du|^2).$$

The solution and its elliptic estimates are obtained by the classical method of sub- and supersolutions; their construction uses in an essential way the hypothesis that the genus of Σ is greater than 1. Indeed, in that case the Euler–Poincaré characteristic of Σ is negative and the Gauss–Bonnet formula permits us to take σ such that $R(\sigma) < 0$; hence $p_3 < 0$.

3.6.2 *Equations for lapse and shift*

Lapse and shift are gauge parameters for which we obtain elliptic equations on each Σ_t as follows.

We impose that the $\Sigma_t's$ have constant (in space) mean curvature. The lapse N then satisfies the linear elliptic equation

$$\Delta N - \alpha N = -e^{2\lambda}\frac{\partial \tau}{\partial t} \tag{3.18}$$

with ($|.|$ pointwise norm in the metric σ, and G if relevant)

$$\alpha \equiv e^{-2\lambda}(\mid h \mid^2 + \mid \dot{u} \mid^2) + \frac{1}{2}e^{2\lambda}\tau^2 \tag{3.19}$$

The equation to be satisfied by the shift ν results from the the expression for h deduced from the definition of k

$$h_{ab} \equiv (2N)^{-1}\left[-\left(\partial_t g_{ab} - \frac{1}{2}g_{ab}g^{cd}\partial_t g_{cd}\right) + \nabla_a \nu_b + \nabla_b \nu_a - g_{ab}\nabla_c \nu^c\right]$$

which implies, if $g_{ab} \equiv e^{2\lambda}\sigma_{ab}$ and if $n_a \equiv e^{-2\lambda}\nu_a$ denotes the covariant components of the shift vector ν in the metric σ (thus $n^a = \nu^a$)

$$h_{ab} \equiv (2N)^{-1}e^{2\lambda}\left[-\left(\partial_t \sigma_{ab} - \frac{1}{2}\sigma_{ab}\sigma^{cd}\partial_t \sigma_{cd}\right) + D_a n_b + D_b n_a - \sigma_{ab}D_c n^c\right]$$

The shift satisfies therefore the non-homogenous linear equation

$$(L_\sigma n)_{ab} \equiv D_a n_b + D_b n_a - \sigma_{ab}D_c n^c = f_{ab} \tag{3.20}$$

with L_σ the conformal Killing operator of σ, and

$$f_{ab} := 2Ne^{-2\lambda}h_{ab} + \partial_t \sigma_{ab} - \frac{1}{2}\sigma_{ab}\sigma^{cd}\partial_t \sigma_{cd} \tag{3.21}$$

The kernel of the conformal Killing operator L_σ depends on the topology of Σ. Recall that the genus of a compact surface is the number of cuts you must inflict to make it simply connected. The genus is zero for S^2, 1 for T^2, and a number greater than 1 for "cushions with more than 1 hole".

Lemma 3.5 *A metric on an oriented 2-surface of genus greater than 1 admits no conformal Killing fields.*

Proof *See results for manifolds with negative curvature in Chapter 7.* □

We have seen in Chapter 7 that the dual of the space of conformal Killing vector fields, the kernel of the operator L_σ, is the space of TT (transverse traceless) tensors, i.e. symmetric 2-tensors T such that

$$\sigma^{ab}T_{ab} = 0, \quad D^a T_{ab} = 0. \tag{3.22}$$

The spaces of TT tensors are the same for two conformal metrics.

3.7 *Teichmüller parameters.*

On a compact 2-dimensional manifold of genus $G > 1$ the space T_{eich} of conformally inequivalent Riemannian metrics, called Teichmüller space, can be identified[16] with M_{-1}/D_0, the quotient of the space of metrics with scalar curvature -1 by the group of diffeomorphisms homotopic to the identity. $M_{-1} \to T_{eich}$ is a trivial fibre bundle whose base can be endowed with the structure of the manifold R^n, with $n = 6G - 6$.

3.7.1 *Gauge choice*

We require the metric σ_t to be in some chosen cross-section $Q \to \psi(Q)$ of the above fibre bundle. Let $Q^I, I = 1, \ldots, n$ be coordinates in T_{eich}; then $\partial\psi/\partial Q^I$ is a known tangent vector to M_{-1} at $\psi(Q)$, that is a symmetric 2-tensor field on Σ, the sum of a transverse traceless tensor field $X_I(Q)$ and of the Lie derivative of a vector field on the manifold $(\Sigma, \psi(Q))$. The tensor fields $X_I(Q), I = 1, \ldots, n$ span the space of transverse traceless tensor fields on $(\Sigma, \psi(Q))$. The matrix with elements

$$\int_\Sigma X_I^{ab} X_{Jab} \mu_{\psi(Q)}$$

is invertible.

3.7.2 *Relation between h and $\partial_t \sigma$*

We have seen that the integrability condition of the shift equation is an equation linking the time derivative of σ_t with the traceless part h_t of the extrinsic curvature. This integrability condition reads, with our definitions

$$\int_{\Sigma_t} f_{ab} X_J^{ab} \mu_{\sigma_t} = 0, \quad J = 1, \ldots, 6G - 6. \tag{3.23}$$

We have

$$\partial_t \sigma_{ab} = \frac{dQ^I}{dt} X_{I,ab} + C_{ab},$$

with C_{ab} a Lie derivative. We denote by P the transverse part of h, with components P^I with respect to the tensor fields $X_{I,Q}$. The integral equation (3.23) gives a linear relation between the tangent vector dQ^I/dt to the curve $t \mapsto Q(t)$ and the tangent vector $P^I(t)$ to T_{eich}

$$X_{IJ} \frac{dQ^I}{dt} + Y_{IJ} P^I + Z_J = 0,$$

with

$$X_{IJ} \equiv \int_{\Sigma_t} X_I^{ab} X_{J,ab} \mu_\sigma, \tag{3.24}$$

$$Y_{IJ} \equiv \int_{\Sigma_t} 2N e^{-2\lambda} X_I^{ab} X_{J,ab} \mu_\sigma, \quad Z_J \equiv \int_{\Sigma_t} 2N e^{-2\lambda} r_{ab} X_J^{ab} \mu_\sigma \tag{3.25}$$

[16] See Fisher, A. and Tromba, A. (1984) *Math. Ann.*, **167**, 118–42, or CB-DM2 V

3.7.3 *ODE for Teichmuller parameters*

An ODE for the function $t \mapsto Q(t)$ is obtained by returning to the full $2+1$ Einstein equations (see details in CB-M2).

We have solved the $2+1$ Einstein constraints and the lapse and shift equations. We can then show, using also the Bianchi identities that the tensor on Σ, ${}^{(3)}R_{ab} - \rho_{ab}$, is a TT tensor. This property is sufficient to prove that ${}^{(3)}R_{ab} - \rho_{ab} = 0$ if Σ is a sphere, but not enough if the genus of Σ is > 0. In our case the necessary and sufficient conditions for the equations already solved to imply also the remaining equations ${}^{(3)}R_{ab} - \rho_{ab} = 0$, are:

$$\int_{\Sigma_t} N({}^{(3)}R_{ab} - \rho_{ab}) X_J^{ab} \mu_{\sigma_t} = 0, \quad \text{for } J = 1, 2, \ldots, 6\ G - 6. \tag{3.26}$$

Straightforward computations using the $2+1$ decomposition show that Equation (3.26) is an ordinary differential system of the form

$$X_{IJ} \frac{dP^I}{dt} + \Phi_J \left(P, \frac{dQ}{dt} \right) = 0$$

where Φ is a polynomial of degree 2 in P and dQ/dt with coefficients depending smoothly on Q and directly but continuously on t through the other unknowns[17].

4 S^1 invariant Einstein universes, Cauchy problem

4.1 *Cauchy data*

The Cauchy data on Σ_{t_0} for the Einstein wave map system are:

1. A C^∞ riemannian metric σ_0 which projects onto a point $Q(t_0)$ of T_{eich} and a C^∞ tensor q_0 which is TT in the metric σ_0.
2. Cauchy data for u and $\dot{u}$ on Σ_{t_0}, i.e.

$$u(t_0, .) = u_0, \qquad \dot{u}(t_0, .) = \dot{u}_0.$$

We say that a pair of scalar functions, $u \equiv (\gamma, \omega)$ or $\dot{u} \equiv (\dot{\gamma}, \dot{\omega})$ belongs to W_s^p if it is so of each of the scalars; W_s^p and $H_s \equiv W_s^2$ are the usual Sobolev spaces of scalar functions on the Riemannian manifold (Σ, σ_0). We suppose that

$$u_0 \in H_2, \qquad \dot{u}_0 \in H_1.$$

From these data one determines the values on Σ_0 of the auxiliary unknown, $h_0 \in W_2^p$, $1 < p < 2$, the conformal factor, lapse and shift λ_0, N_0, $\nu_0 \in W_3^p$.

One deduces then the usual Cauchy data for the wave map by

$$(\partial_t u)_0 = e^{-2\lambda_0} N_0 \dot{u}_0 + \nu_0^a \partial_a u_0 \tag{4.1}$$

It holds that

$$(\partial_t u)_0 \in H_1. \tag{4.2}$$

[17] See details in CB-M2.

4.2 *Construction of A when F is known*

A solution of the Einstein wave map system determines ω, hence the curvature F, but the unknown in the four-dimensional problem is the 1-form A, representative of the connection 1-form θ. This 1-form can be constructed when the curvature F is known if the following integrability condition is satisfied (see M1, CB3)

$$\int_{\Sigma_t} F = \frac{1}{2}\int_{\Sigma_t} F_{ab}dx^a \wedge dx^b = -\int_{\Sigma_t} e^{-4\gamma}N^{-1}\partial_0\omega\mu_g = 2\pi n. \tag{4.3}$$

where n is the Chern number of the bundle over $\Sigma\times R$. This condition is conserved under time evolution by a solution of the wave map equation.

In fact the value $n = 0$ (trivial bundle) is the only one compatible with the smallness assumptions on the energy that we will make. The 1-form $A \equiv A_\alpha dx^\alpha$ such that $dA = F$ is then defined globally[18] on $\Sigma \times R$, uniquely up to the value of A_t.

4.3 *Local in time existence theorem*

A local existence theorem can be proved by the standard iterative method of solving successively hyperbolic equations (the wave map for a given metric), elliptic equations (constraints and gauge equations for a given wave map) and ODE (for the Teïchmuller parameters). The order of necessary derivatives matches, and the approximations converge for a small enough time interval (see details of proof in CB-M1). We state the theorem.

Theorem 4.1 *The Cauchy problem with the above data for the Einstein equations with S^1 isometry group has, if $T - t_0$ is small enough, a solution with $(\gamma, \omega) \in C^0([t_0, T), H_2)$, $(\dot\gamma, \dot\omega) \in C^1([t_0, T), H_1)$; $\lambda, N, \nu \in C^0([t_0, T), W_3^p)$ $\cap C^1([t_0, T), W_2^p)$, $1 < p < 2$ and $N > 0$ while $\sigma \in C^1([t_0, T), C^\infty)$ with σ_t uniformly equivalent to σ_0. This solution is unique up to t parametrization of τ, choice of A_t, and choice of a cross-section of M_{-1} over T_{eich}.*

4.4 *Global existence theorem*

4.4.1 *General scheme*

The global existence theorem relies in an essential way on the expansion of the universe. We look for a metric with constant mean curvature slicing and an expanding universe; that is the mean curvature τ starts negative and increases. The universe attains a moment of maximum expansion if the metric is defined up to $\tau = 0$. We choose the time parameter t by requiring that

$$t = -\frac{1}{\tau}. \tag{4.4}$$

Then t increases from $t_0 > 0$ to infinity when τ increases from $\tau_0 < 0$ to zero.

[18] See details in CB3.

It results from the local existence theorem and the standard arguments that the solution of the Einstein equations exists for $t \in [t_0, \infty)$ if the norms $||u(t,.)||_{H_2}$, $||\partial_t \gamma(t,.), \partial_t \omega(t,.)||_{H_1}$ on (Σ, σ_t) do not blow up for any finite t, the metric σ_t remaining smooth and uniformly equivalent to σ_0, i.e. the curve $t \mapsto Q(t)$ remaining in a compact subset of T_{eich}.

The proof requires energy and higher energy estimates for the wave map when the metric of the source is estimated via elliptic estimates in Sobolev spaces on (Σ, σ_t). These elliptic estimates themselves depend on wave map estimates. The proof of the equivalence of norms in Sobolev norms in various metrics σ_t depends on the other estimates through the coefficients of the ODE satisfied by Q. These estimates require the introduction of corrected energies. The various inequalities are obtained in assuming *a priori* bounds. The final estimate is of the form of a simple bootstrap argument.

4.4.2 *First energy estimate*

Inspired by the Hamiltonian formulation of General Relativity, one defines the first energy not only as the energy of the wave map but by the L^2 norm of the traceless part h of the extrinsic curvature. We set

$$E(t) \equiv \int_{\Sigma_t} \left(|\partial u|_g^2 + |u'|^2 + \frac{1}{2}|h|_g^2 \right) \mu_g,$$

where $|.|_g$ denotes the point wise norm in the 2-metric g and also, when the wave map is concerned, the Poincaré plane metric G.

The Hamiltonian constraint

$$R(g) + \frac{\tau^2}{2} = |\partial u|_g^2 + |u'|^2 + \frac{1}{2}|h|_g^2$$

together with the Gauss–Bonnet formula (χ the Euler Poincaré constant)

$$\int_{\Sigma_t} R(g)\mu_g = 4\pi\chi \tag{4.5}$$

give, without using the wave map equation, that:

$$\frac{dE(t)}{dt} = \frac{1}{2}\tau \int_{\Sigma_t} N \left(|u'|^2 + \frac{1}{2}|h|_g^2 \right) \mu_g \leq 0. \tag{4.6}$$

We deduce from this equality that $E(t)$ is a non-increasing function of t since $\tau < 0$, but we do not obtain its decay due to the absence of $|Du|^2$ from the right-hand side.

4.4.3 *First elliptic estimates*

The definition of $E(t)$ implies

$$||h||_{L^2(\sigma)} \leq e^{2\lambda_M} E(t)$$

while under the hypothesis $R(\sigma) = -1$ the maximum principle applied to (3.17) implies

$$e^{2\lambda} \geq 2\tau^{-2} \tag{4.7}$$

and, with the parameter choice $\tau = -\frac{1}{t}$, it implies, applied to (3.18)

$$0 < N \leq 2. \tag{4.8}$$

Further elliptic estimates require bounds of $\partial u.\partial u$ and $\partial u.u'$ in $W_1^p(\sigma)$, $1 < p < 2$, which are obtained in terms of the second energy of the wave map.

4.4.4 *Second energy*

We denote by $\hat{\nabla}$ and $\hat{\partial}_0$ covariant derivatives for mappings (Σ, g) or (R, dt^2) into the Poincaré plane (R^2, G). We define the second energy by

$$E^{(1)}(t) := \int_{\Sigma_t} (\mid \hat{\Delta}_g u \mid^2 + \mid \hat{\nabla} u' \mid^2)\mu_g \tag{4.9}$$

We use the notations

$$\varepsilon^2 := E(t), \quad \varepsilon_1^2 := \tau^{-2} E_1(t). \tag{4.10}$$

We find after long computations using elliptic estimates applied to the constraints and lapse equations that

$$1 \leq \frac{1}{\sqrt{2}} |\tau| e^{\lambda} \leq 1 + C_{E,\sigma}(\varepsilon + \varepsilon_1), \tag{4.11}$$

$$0 \leq 2 - N \leq C_{E,\sigma}(\varepsilon^2 + \varepsilon\varepsilon_1) \tag{4.12}$$

where we denote by $C_{\sigma,E}$ numbers depending only on *a priori* bound of ε and ε_1, and on the domain of σ in *Teich* supposed to be compact.

We use these bounds and the equation satisfied by the wave map $(\Sigma \times R, {}^{(3)}g) \to (R^2, G)$ which reads in our notations:

$$-N^{-1}\hat{\partial}_0 \partial_0 u + g^{ab}\hat{\nabla}_a(N\partial_b u) + Nu' = 0. \tag{4.13}$$

We find, also because the target of the wave map has negative constant curvature, that

$$\frac{dE^{(1)}}{dt} - 2\tau E^{(1)} = \tau \int_{\Sigma_t} N \mid Du' \mid^2 \mu_g + Z \leq Z, \quad \tau < 0) \tag{4.14}$$

where Z satisfies an inequality of the form

$$|Z| \leq |\tau|^3 C_{\sigma,E}(\varepsilon + \varepsilon_1)^3. \tag{4.15}$$

The higher than one power in the energies in the positive part of the right-hand side is a hopeful sign for the boundedness of these energies. However, the

inequalities (4.14) and (4.15), are not sufficient to prove the bound of $\varepsilon + \varepsilon_1$ and the fact that σ_t projects on a fixed compact subset of T_{eich}, using the ODE satisfied by Q. The proof of the boundedness of Q requires decay of $\varepsilon + \varepsilon_1$.

4.4.5 *Corrected energies*

To obtain the decay property one introduces corrected energies and exploits the negative (non-definite) terms in the energies inequalities. We define the first corrected energy to be

$$E_\alpha(t) := E(t) - \alpha\tau \int_{\Sigma_t} (u - \bar{u}).u' \mu_g \tag{4.16}$$

where $\bar{u}$ denotes the mean value of u on Σ in the metric σ; that is,

$$\bar{u} = \frac{1}{Vol_{\sigma_t}} \int_{\Sigma_t} u \mu_\sigma.$$

We define the second corrected energy by

$$E^{(1)}_\alpha(t) = E^{(1)}(t) + \alpha\tau \int_{\Sigma_t} \hat{\Delta}_g u.u' \mu_g \tag{4.17}$$

The use of elliptic estimates leads to

$$\frac{dE_\alpha}{dt} - k\tau E_\alpha \leq |\tau| C_{\sigma,E}(\varepsilon + \varepsilon_1)^3, \tag{4.18}$$

and

$$\frac{dE^{(1)}_\alpha}{dt} - (2+k)\tau E^{(1)}_\alpha + |\tau|^3 C_{\sigma,E}(\varepsilon + \varepsilon_1)^3. \tag{4.19}$$

We denote by Λ_σ the first positive eigenvalue of $-\Delta_\sigma$ and we prove that $E_\alpha + \tau^{-2} E^{(1)}_\alpha$ is equivalent to the total original energy $E_{tot} := \varepsilon^2 + \varepsilon_1^2$ under the following conditions:

$$\begin{gathered} \alpha = \frac{1}{4}, \quad k = 1 \quad \text{if} \quad \Lambda_\sigma > \frac{1}{8} \\ \alpha < \frac{4}{8 + \Lambda_\sigma^{-1}}, \quad 0 < k < 1, \quad \text{if} \quad \Lambda_\sigma \leq \frac{1}{8}. \end{gathered} \tag{4.20}$$

Lemma 4.2 *Given an* a priori *bound of the total original energy* $E_{tot} := \varepsilon^2 + \varepsilon_1^2$ *and of the domain of* Q *in Teichmuller space, the total energy and* Q *satisfy inequalities of the form*

$$E_{tot}(t) \leq t^{-k} M_1 E_{tot}(t_0), \quad \textit{for some } k > 0, \tag{4.21}$$

$$|Q(t) - Q(t_0)| \leq M_2 E_{tot}(t_0)(t_0). \tag{4.22}$$

with M_1 *and* M_2 *numbers depending only on the given* a priori *bounds*

Proof The inequality for E_{tot} results from the differential inequalities satisfied by the corrected energy and the equivalence with the total original energy. The inequality for Q results from the differential equation it satisfies and the decay of the total energy. □

4.5 *Future complete existence*

Theorem 4.3 *Let $(\sigma_0, q_0) \in C^\infty(\Sigma_0)$ and $(u_0, \dot{u}_0) \in H_2(\Sigma_0, \sigma_0) \times H_1(\Sigma_0, \sigma_0)$ be initial data on the compact manifold Σ_0, Genus$(\Sigma_0) > 1$, satisfying the Chern integrability condition. There exists a number $\eta > 0$ such that, if $E_{tot}(t_0) < \eta$, the vacuum four-dimensional Einstein equations have a solution on $\Sigma \times S^1 \times [t_0, \infty)$, $t = -\tau^{-1}$, with initial values determined by $\sigma_0, q_0, u_0, \dot{u}_0$. This space time is globally hyperbolic and future timelike and null complete.*

Proof

1. We use what is called a bootstrap argument to prove an *a priori* bound of $E_{tot}(t)$.

Suppose the given *a priori* bounds corresponding to M_1 and M_2 are

$$E_{tot}(t) \leq C_1, \quad |Q(t) - Q(t_0)| \leq C_2.$$

Then the inequalities (4.21) and (4.22) imply that if $E_{tot}(t_0)$ is small enough then the pair $(E_{tot}(t), Q(t) - Q(t_0))$ belongs to the open set

$$U_1 := \{E_{tot}(t) < C_1, \quad |Q(t) - Q(t_0)| < C_2\}.$$

The two inequalities show that the pair $(E_{tot}(t), Q(t) - Q(t_0))$ belongs either to the open set U_1 or to the open set

$$U_2 := \{E_{tot}(t) > C_1, \quad |Q(t) - Q(t_0)| > C_2\}.$$

The two open sets U_1 and U_2 are disjoint. The pair $(E_{tot}(t), Q(t) - Q(t_0))$ depends continuously on t,it stays in one of the open sets, in $U(1)$ if the initial data are ssuch that $E_{tot}(t_0) < C_1$.

2. One deduces the existence of the solution on $\Sigma \times [t_0, \infty)$, and the existence for an infinite proper time along the lines $\{x\} \times R$ after estimating the usual H_2 norms in terms of the geometrically defined second energy. This estimate depends[19] on the fact that the Riemann curvature of the target metric is constant and negative.

The global hyperbolicity and completeness is a particular case of a theorem proved by CB and Cotsakis (see Chapter 12). □

4.6 *Einstein–Maxwell–Higgs system*

The global existence theorem extends to the Einstein–Maxwell–Higgs system with a spatial S^1 isometry group, space sections a principal S^1 fibre bundle with

[19] See details in CB3.

basis a 2-surface of genus greater than 1, and electromagnetic field orthogonal to the Killing field.

The proof[20] uses the fact that such a system can be formulated[21] as a three-dimensional Einstein system with source a wave map in a manifold (R^4, G) with G a Riemannian metric of constant negative curvature.

4.7 *Conclusion*

Besides the fact that the $U(1)$ symmetric case is not included in the no symmetry case treated by Andersson and Moncrief[22] (see next section, an interest of the $U(1)$ case is that in our problem the number of effective spatial dimensions is two, and also that there is no known "physical" reason why large data solutions should develop singularities in the direction of cosmological expansion. Black hole formation seems to be suppressed by the topological character of the assumed Killing symmetry (which is of translational rather than rotational type and excludes the appearance of an axis of symmetry) and the Big Bang singularity is avoided by considering the future evolution from an initially expanding Cauchy hypersurface. Any possible big crunch is excluded by our requirement that the spatial manifold M is of negative Yamabe type (which is true of all circle bundles over higher genus manifolds). Such manifolds are incompatible (in the vacuum and electrovacuum cases for example) with the development of a maximal hypersurface which would be a necessary prelude to the "recollapse" of an expanding universe towards a hypothetical big crunch singularity. At a maximal hypersurface the scalar curvature of M would have to be everywhere positive – an impossibility on any manifold of negative Yamabe type. Thus it is conceivable that for large data future global existence holds for our problem. Up to now the only large data global results require simplifying assumptions so stringent that they effectively reduce the number of spatial dimensions to one (e.g. Gowdy models and their generalizations, plane symmetric gravitational waves, spherically symmetric matter coupled with gravity) or zero (e.g. Bianchi models, $2+1$ gravity). Unfortunately we have at present no way of proving this global existence, even in the polarized case for which the wave map equation reduces to a wave equation for a scalar function, because the reduced field equations are nonlocal in character. The "background" spacetime on which the scalar field evolves is not given *a priori* but is instead a certain functional (obtained by solution of elliptic equations) of the evolving field (and the Teichmüller parameters) itself. In the unpolarized case, the problem of global existence of strong solutions for wave maps on a fixed background in $2+1$ dimensions is still unsolved. However, there is a proof by Müller and Struwe[23] of the global existence of a weak solution (with

[20] Choquet-Bruhat, Y. (2005) Proc. Greek Relativity Meeting, Lesbos, *J. Phys. Conf. Ser.*, **8**, 1–12, gr-qc 0501052.

[21] Moncrief, V. (1990) *Class. Quant. Grav.*, **7**, 320–52.

[22] These comments are part of the introduction written by V. Moncrief for the paper CB3.

[23] S. Muller and M. Struwe (1996) *Top. Meth. Non-Lin. Anal.*, **7**(2), 245–61.

no uniqueness) for wave maps from Minkowski spacetime, and, as yet unpublished, work of T. Tao on global existence of strong solutions. Any progress on the large data global existence, even of weak solutions, for the $U(1)$-symmetric problem would represent a "quantum jump" forward in our understanding of long time existence problems for Einstein's equations.

5 Andersson–Moncrief theorem

The small data future global existence theorem of Andersson and Moncrief for solutions of vacuum Einstein's equations[24], that we briefly present, makes no symmetry assumption on the spacetime metric, but treats a different class of spatial 3-manifolds than the manifolds considered in the previous section. Their result concerns the non-linear stability of a flat manifold with compact spacelike sections which they call a hyperbolic cone spacetime, defined as follows.

Definition 5.1 *A hyperbolic cone spacetime is the manifold $(0,\infty)\times M$, with M a compact manifold, endowed with the Lorentzian metric*

$$-d\rho^2+\rho^2\gamma,$$

where γ is a Riemannian metric, independent of ρ, of constant Riemannian[25] *curvature -1.*

Remark that the homogeneous 3-space (M,γ), with M of negative Riemannian curvature and compact, is necessarily multiply connected (see Chapter 5).

Lemma 5.2 *A hyperbolic cone spacetime is flat.*

Proof The space metric $\rho^2\gamma_{ij}$ and the metric γ_{ij} differ by a constant (in space) conformal factor. The covariant derivatives in these two metrics are the same, their Riemann tensors (in covariant components) differ by a factor ρ^2. The extrinsic curvature of the space slices is

$$K_{ij}=-\rho\gamma_{ij}\equiv-\frac{1}{\rho}g_{ij}.$$

The decomposition formulae of Chapter 6 with the lapse $N=1$ and the shift $\beta=0$ give immediately that the Riemann tensor of the spacetime is identically zero. □

Andersson and Moncrief prove a global existence theorem of vacuum Einsteinian spacetimes for small data, that is near the hyperbolic cone initial data, by using a hyperbolic-elliptic system. Their spacetimes (V,g) have the usual $3+1$ decomposition,

$$g:=-N^2dt^2+g_{ij}(dx^i+\beta^i dt)(dx^j+\beta^j dt).$$

[24] Andersson, L. and Moncrief, V. (2004) in *The Einstein Equations and the Large Scale Behavior of Spacetime* (eds. P. Chruściel and H. Friedrich) Birkhaüser, pp. 299–330.

[25] Equivalently, sectional. We avoid the name hyperbolic often given to such Riemannian metrics to avoid confusion with hyperbolic signature.

The hyperbolic system is composed of Bianchi equations for the Weyl tensor. The elliptic system determines the t-dependent space metric $\bar{g}$ and extrinsic curvature K when the Weyl tensor is known, modulo gauge choices which give elliptic equations for the lapse and the shift (see Chapter 8).

5.1 *CMC gauge, elliptic system for N and K*

5.1.1 *CMC gauge*

CMC time gauge means that the extrinsic curvature τ of each space slice M_t is constant. In CMC gauge, the lapse N satisfies on each M_t the elliptic equation

$$\Delta_{\bar{g}} N - |K|^2 N = -\partial_t \tau.$$

In the case of the hyperbolic cone this condition means a choice of time t such that $\tau \equiv \rho^{-2}\gamma^{ij}K_{ij} \equiv -3\rho^{-1}$ depends only on t. Andersson and Moncrief choose $t = \tau$. In this gauge the metric reads

$$-\frac{9}{\tau} d\tau^2 + \frac{9}{\tau^2}\gamma_{ij} dx^i dx^j.$$

As τ, the mean curvature of the space slices, varies from $-\infty$ to zero their volume increases from zero to infinity.

5.1.2 *Elliptic system for K*

We have seen in Section VIII.6.5 that the identity of the $3+1$ decomposition,

$$\bar{\nabla}_j K_{ki} - \bar{\nabla}_k K_{ji} = N^{-1} R_{0i,jk}, \tag{5.1}$$

together with the CMC gauge condition and the momentum constraint

$$\partial_i Tr K = 0, \qquad \bar{\nabla}_j K_i^j = 0, \tag{5.2}$$

implies that K satisfies a linear self-adjoint second-order elliptic system. The general theory of such inhomogeneous linear systems on compact manifolds says that they are solvable, with solutions satisfying Sobolev estimates if the kernel of the adjoint operator reduces to zero[26]. Symmetric 2-tensors satisfying the equations (5.1) with zero right-hand side are called **Codazzi tensors**. Compact Riemannian manifolds which do not admit non-identically zero traceless Codazzi tensors are called **rigid**. The rigidity of the manifold $(M, \bar{g})$ is a sufficient condition for the elliptic system satisfied by K to be well posed.

5.2 *SH gauge, elliptic system for $\bar{g}$ and β*

5.2.1 *Spatial harmonic (SH) gauge. Equation for $\bar{g}$*

The spatial harmonic gauge is the equivalent for Riemannian manifolds to the wave gauge for Lorentzian manifolds (see Chapter 6).

[26] We have seen that this kernel is indeed zero for compact 3-manifolds with constant negative curvature.

The space metric $\bar{g}_t$ is said to be in SH gauge on$M$ (gauge introduced by Andersson and Moncrief[27], discussed in Chapter 8) if the mapping $(M, g_t) \to (M, \gamma)$ is a harmonic diffeomorphism. This condition is trivially satisfied for the metric γ itself. It can be imposed on the perturbed metric $\bar{g}_t$, since the target metric γ has constant negative curvature, the required diffeomorphism exists. In SH gauge the Ricci tensor $\bar{R}_{ij}$ is an elliptic operator for $\bar{g}$. The identity linking R_{ij} with $\bar{R}_{ij}$ and K gives an elliptic second-order differential system for $\bar{g}$ when R_{ij} is known.

5.2.2 *Shift equation*

An elliptic equation for the shift results from the Bianchi identities.

5.3 *The Bianchi equations*

Andersson and Moncrief choose as evolution system the Bianchi equations satisfied by the Weyl tensor, identical in vacuum to the Riemann tensor. They estimate the Weyl tensor and its first derivatives via the Bel–Robinson energy, as do Christodoulou and Klainerman, but they do not need to introduce its Lie derivatives with respect to privileged vector fields: behaviour at space infinity does not cause problems in a compact space.

5.3.1 *Bel–Robinson energies*

The local existence theorem requires, in this case with no symmetries, that the space metrics be in H_3, hence estimates of the Bel–Robinson energy of the Weyl tensor and of its first derivatives.

AM compute in a standard way the differential equation satisfied by the **Bel–Robinson energy** $\mathcal{B}(t)$ of the Weyl tensor at time t, defined from the **Bel tensor** Q and the unit normal to the space slices, a timelike vector which they denote T:

$$\mathcal{B}(t) := \int_{M_t} Q(T,T,T,T) d\mu.$$

They prove the following lemma, which is a formulation adapted to the problem at hand of an equality given in Chapter 8.

Lemma 5.3 *In a vacuum spacetime with a constant mean curvature slicing $\tau = TrK$ the Bel–Robinson energy $\mathcal{B}$ satisfies the differential equality*

$$\partial_\tau \mathcal{B}(\tau) = \frac{3}{\tau}\mathcal{B}(\tau) - \int_{M_\tau} N Q_{\alpha\beta\gamma\delta} \hat{\pi}^{\alpha\beta} T^\gamma T^\delta d\mu_{\bar{g}} + \tau \int_{M_\tau} \hat{N} Q(T,T,T,T) d\mu_{\bar{g}},$$

where $\hat{\pi}$ is the "trace free" part of the deformation tensor

$$\hat{\pi}^{\alpha\beta} := \pi^{\alpha\beta} + \frac{\tau}{3}(g^{\alpha\beta} + T^\alpha T^\beta), \quad \pi_{\alpha\beta} := \frac{1}{2}(\mathcal{L}_T g)_{\alpha\beta}$$

[27] Andersson, L. and Moncrief, V. (2001) *Ann. H. Poincaré*, **4**(1), 1–34.

and $\hat{N}$ *the "perturbed" part of the lapse*

$$\hat{N} := N - \frac{3}{\tau^2}.$$

Andersson and Moncrief denote by Q_1 the Bel tensor of $\nabla_T W$, which is also a Weyl field; that is, it has the same symmetries and antisymmetries as the Weyl (or Riemann) tensor and which is traceless. Denoting by $\mathcal{B}_1(t)$ the Bel–Robinson energy defined through Q_1. AM obtain, by long calculations, a differential equality for Q_1.

Lemma 5.4 *In a vacuum spacetime with a constant mean curvature slicing the Bel–Robinson energy* $\mathcal{B}_1$ *of the tensor* $\nabla_T W$ *satisfies a differential equality of the form*

$$\partial_\tau \mathcal{B}_1(\tau) = \frac{5}{\tau}\mathcal{B}_1(\tau) - 3\int_{M_\tau} NQ_{1,\alpha\beta\gamma\delta}\hat{\pi}^{\alpha\beta}T^\gamma T^\delta d\mu_{\bar{g}} + \frac{5\tau}{3}\int_{M_\tau} \hat{N}Q_1(T,T,T,T)d\mu_{\bar{g}} - 2\int_{M_\tau} \mathcal{G}_1(W,\nabla W)d\mu_{\bar{g}}.$$

The explicit expression of the term $\mathcal{G}_1$ can be found in the work of Andersson and Moncrief. The important property is that it is cubic in $\hat{\pi}$, W, and ∇W, and hence can be treated as a perturbation term in case of small data, i.e. near the flat hyperbolic cone initial data.

5.4 *Existence theorems*

5.4.1 *Local existence theorem*

A detailed proof of a local existence theorem for their elliptic-hyperbolic theorem has been given by Andersson and Moncrief[28].

Theorem 5.5 *Let* (M,γ) *be a compact smooth Riemannian manifold with* γ *a metric of constant Riemannian curvature* -1. *Let* $(M,\bar{g}_0,K_0)$ *be an initial data set with* $\bar{g}_0$ *in SH gauge with respect to* γ, $Tr_{\bar{g}_0}K_0 = 0$ *and* $(\bar{g}_0,K_0) \in H_3 \times H_2$ *on* (M,γ). *If these data satisfy the Einstein vacuum constraints they admit a vacuum Einsteinian development* (V,g) *with slicing in CMCSH gauge, the* $(H_3 \times H_2)(M,\gamma)$ *norms of* $(\bar{g}_t,K_t)$ *depending continuously on the norms of the initial data.*

5.4.2 *Global estimates*

The global existence theorem results, as usual, from *a priori* bounds on the norms at time t of the same kind than norms appearing for the initial data, and of the continuous dependence of the interval of existence from these norms.

The elliptic estimates would naturally be made with Sobolev norms on the Riemannian manifolds $(M_\tau,\bar{g}_\tau)$, which depend on τ. Fortunately in the problem considered by AM, they are able to introduce a scaling such that they have

[28] Andersson, L. and Moncrief, V. (2001).

universal norms without the difficulties we met with controlling norms in the conformal metric σ_t through estimates in the Teichmüller space.

AM define what they call the scale-free BR energies, by

$$\tilde{\mathcal{B}} = \lambda^{-1}\mathcal{B}, \quad \tilde{\mathcal{B}}_1 = \lambda^{-3}\mathcal{B}_1, \quad \text{with} \quad \lambda := -\frac{\tau}{3},$$

and the total scale free BR energy, a function of the time τ, by

$$\mathcal{E} = \tilde{\mathcal{B}} + \tilde{\mathcal{B}}_1.$$

Andersson and Moncrief introduce the scaled variables

$$\tilde{g}_{ij} := \lambda^2 \bar{g}_{ij}, \quad \tilde{N} := \lambda^2 N, \quad \text{then} \quad \tilde{T}^\alpha \equiv \lambda^{-1} T \quad \text{and} \quad \mu_{\tilde{g}} = \lambda^3 \mu_{\bar{g}};$$

They use the conformal invariance of the Weyl tensor in 3 covariant 1 contravariant form, and the previous equations to show that $\mathcal{E}(\tau)$ is a smooth functional of the pair $(\bar{g}_\tau, K_\tau) \in H_3 \times H_2$. At the flat hyperbolic cone, this functional and its first Frechet derivative are zero, since it is so of the Weyl tensor and since the BR energy is quadratic in it. Andersson and Moncrief compute the second derivative of $\mathcal{E}$, also called the Hessian, at the flat hyperbolic cone, quadratic form in the pair $(h, p) := (\bar{g}_\tau - \frac{9}{\tau^2}\gamma, K_\tau - \frac{3}{\tau}\gamma)$. They prove the following lemma.

Lemma 5.6 *The Hessian of $\mathcal{E}$ at $(\frac{9}{\tau^2}\gamma, \frac{3}{\tau}\gamma)$ satisfies the inequality*

$$(Hess\mathcal{E})\left(\frac{9}{\tau^2}\gamma, \frac{3}{\tau}\gamma\right)(h,p)(h,p) \geq C(||h||_{H_3} + ||p||_{H_2})$$

if and only M is rigid. The positive constant C depends only on the topology of M.

From this lower bound and the Taylor formula Andersson and Moncrief deduce a bound by the total scale free energy of the $H_3 \times H_2$ norms of $(\frac{\tau^2}{9}\bar{g}_\tau - \gamma, \frac{\tau}{3}K_\tau - \gamma)$.

They write differential inequalities satisfied by the total rescaled BR energy, replacing for convenience the CMC time τ by what they call logarithmic time,

$$\sigma := -\ln(-\tau), \quad \text{then} \quad \sigma \to +\infty \quad \text{when} \quad \tau \to 0.$$

and prove the following lemma under a previous hypothesis on (M, γ).

Lemma 5.7 *If the initial data $(\bar{g}_0, K_0)$ are sufficiently near in $H_3 \times H_2$ norm from the hyperbolic cone initial data, then the total scaled BR energy is so that*

$$\partial_\sigma \mathcal{E} \leq -2(2 - 2C\mathcal{E}^{\frac{1}{2}})\mathcal{E}.$$

5.5 *Global existence theorem*

We state the Andersson–Moncrief theorem (almost) in their own terms, after giving their definition of smallness of an initial data set.

Definition 5.8 *An initial data set* $(\bar{g}_0, K_0)$ *at CMC time* τ_0 *satisfies the* ε *smallness condition if*

$$\left|\left|\frac{\tau_0^2}{9}\bar{g}_0 - \gamma\right|\right|_{H_3} + \left|\left|\frac{\tau}{3}K_0 - \gamma\right|\right|_{H_2} < \varepsilon.$$

Theorem 5.9 *Assume that* M *is rigid. Let the initial data set* $(\bar{g}_0, K_0)$ *at CMC time* $\tau_0 < 0$ *satisfy the vacuum Einstein constraints and the* ε *smallness condition. If* ε *is small enough there exists a vacuum Einsteinian development in CMCSH gauge with* $T_- < \tau < 0$, $\tau_0 \in (-\infty, 0)$. *This development is globally hyperbolic and causally geodesically complete in the expanding direction.*

6 Einstein non-linear scalar field system

As we mentioned in Chapter 5, recent observations favour accelerated expansion of the universe. The physical mechanism of this acceleration is still unclear, but a possible mathematical model is through the existence of a scalar field. A particular case of this involves a positive cosmological constant, as in the case of the De Sitter spacetime which has exponential expansion and was proved long ago by Friedrich to enjoy non-linear stability. The more general case of Einsteinian spacetimes with source a scalar field with a non-linear potential is a subject of active study. Remarkable results for such Einsteinian spacetimes have been obtained particularly by Rendall[29] on the asymptotic behaviour and by Ringström[30] on the non-linear stability in the case of locally homogeneous compact space sections. The proofs rely on both geometry and analysis.

[29] Rendall, A. (2004) *Class. Quant. Grav.*, **21**, 2445–54, and references therein,
[30] Ringström, H. to appear in *Invent. Math.*

APPENDIX I

SOBOLEV SPACES ON RIEMANNIAN MANIFOLDS

Sobolev spaces on Riemannian manifolds have been extensively studied[1]. To remain comparatively simple we consider in this book only Sobolev spaces of integral order[2].

1 Definitions

Let (M, e) be a smooth Riemannian manifold. We denote by ∂ the covariant derivative and by μ_e the volume element in the metric e.

Definition 1.1

1. *The Sobolev space W_s^p is the space of functions, or tensor fields of some given type, on (M, e) with L^p integrable generalized derivatives of order $\leq s$ in the metric e. It is a Banach space with norm*

$$\| f \|_{W_s^p} := \left\{ \int_M \sum_{0 \leq k \leq s} | \partial^k f |^p \, \mu_e \right\}^{1/p}, \quad 1 \leq p < \infty \tag{1.1}$$

with $| \, |$ the pointwise norm of tensors in the metric e.

2. *The Sobolev (Banach) spaces $\overset{0}{W}{}_s^p \subset W_s^p$ on (M, e) are closures with respect to the norms (1.1) of spaces $\mathcal{D}$ of C^∞ functions, or tensor fields of some given type, with compact support in M.*

Proposition 1.2

1. *The space $\overset{0}{W}{}_0^p$ is identical with the space $W_0^p := L^p(M, e)$.*

2. *The space $\overset{0}{W}{}_1^p$ coincides with the space W_1^p when (M, e) has a non-zero injectivity radius, hence when (M, e) is complete. It is true in particular if M is a compact manifold (without boundary). On the other hand if (M, e) can be*

[1] See for instance Aubin, T. (1982) *Non-Linear Analysis and Monge–Ampère Equations*, Springer, or CB-DM2 VI, 2, 3, 6 and references therein.

[2] Sobolev spaces of fractional order are defined on Euclidean spaces through Fourier analysis. Recent papers have extended definitions and properties by localized study (see Klainerman and Rodnianski arXiv.math.AP0109174 and references therein). Sobolev spaces with negative s are defined by duality.

identified with an open relatively compact subset, with differentiable boundary ∂M*, of another smooth Riemannian manifold then functions in* $\overset{0}{W}{}_1^p$ *have a zero trace on* ∂M*: they constitute only a subset of* W_1^p*.*

3. *The space* $\overset{0}{W}{}_s^p$*,* $s \geq 2$*, coincides with* W_s^p *if* (M, e) *has a non-zero injectivity radius and a curvature tensor uniformly bounded as well as its derivatives up to order* $s - 2$*.*

2 Embedding and multiplication properties

2.1 *Open subsets of* R^n

One says that an open set U of R^n endowed with the Euclidean metric satisfies the **cone property** if there exists $\alpha > 0$ and $h > 0$ such that at each point $x \in U$ one can draw an axially symmetric cone of vertex x, angle α and height h contained in U. The spaces W_s^p on an open set $U \subset R^n$ with the cone property satisfy the following embedding and multiplication properties.

Proposition 2.1 ***Embedding properties***

1. *If* $s > \frac{n}{p}$ *then the space* W_s^p *is continuously imbedded into* $\bar{C}^0$ *(continuous and bounded functions); it holds that, with* $C(U)$ *a number depending only on* U

$$Sup_M |u| \leq C(U)||u||_{W_s^p}. \tag{2.1}$$

Also $W_n^1 \subset \bar{C}^0$*.*

2. *If* $s = \frac{n}{p}$*, the space* W_s^p *is continuously embedded in* L^q *for all* q *such that*

$$p \leq q < \infty. \tag{2.2}$$

3. *If* $s < \frac{n}{p}$*, the space* W_s^p *is continuously embedded in* L^q *for all* q *such that*

$$p \leq q \leq \frac{np}{n - sp}. \tag{2.3}$$

In cases 2 and 3 it holds that

$$||u||_{L^q} \leq C(U)||u||_{W_s^p}. \tag{2.4}$$

Corollary 2.2 *An easy induction shows the continuous embeddings:*

$$W_{s+m}^p \subset \bar{C}^m \quad \textit{if} \quad s > \frac{n}{p}. \tag{2.5}$$

$$W_{s+m}^p \subset W_m^q, \quad \textit{if} \quad s < \frac{n}{p}, \quad p \leq q \leq \frac{np}{n - sp}. \tag{2.6}$$

Proposition 2.3 *The spaces W^p_s admit in U the continuous* ***multiplication property****:*

$$W^p_{s_1} \times W^p_{s_2} \to W^p_s, \quad by \quad (u,v) \mapsto u \otimes v,$$
$$||u \otimes v||_s \leq C(U,e)||u||_{W^p_{s_1}}||u||_{W^p_{s_2}} \quad if \quad s_1 + s_2 > s + \frac{n}{p}, \quad s_1, s_2 \geq s$$

In particular W^p_s is an algebra if $s > \frac{n}{p}$.

The following theorem is particularly useful in elliptic estimates.

Theorem 2.4 *(Gagliardo–Nirenberg interpolation inequality) When $M = R^n$ or M is a compact manifold with C^m boundary*[3]*, and when $f \in W^p_m \cap L^r$, $1 \leq p \leq q \leq r \leq \infty$, it holds that*

$$||\partial^j f||_{L^q} \leq C||f||^{\lambda}_{W^p_m}||f||^{1-\lambda}_{L^r}, \quad if \quad \frac{j}{m} \leq \lambda \leq 1, \tag{2.7}$$

with[4]

$$\lambda = \left(\frac{j}{n} + \frac{1}{r} - \frac{1}{q}\right)\left(\frac{m}{n} + \frac{1}{r} - \frac{1}{p}\right)^{-1}. \tag{2.8}$$

For any $\varepsilon > 0$, there is a number C_ε such that, for all $u \in W^p_{m,\delta}$, $1 \leq p \leq q \leq \infty$, $j < m$

$$||\partial^j u||_{L^q} \leq C\{\varepsilon||u||_{W^p_m} + C_\varepsilon||u||_{W^p_0}\}. \tag{2.9}$$

2.2 *Riemannian manifolds*

We say that the smooth Riemannian manifold (M,e) is **Sobolev regular** if the spaces W^p_s defined on (M,e) satisfy the embedding and multiplication properties of the above proposition, the numbers $C(U)$ being replaced by numbers $C(M,e)$, called Sobolev constants, depending only on M and e. They then satisfy the interpolation inequality.

Theorem 2.5 *Complete Riemannian manifolds, or Riemannian manifolds with Lipshitzian boundary, are Sobolev regular.*

Proof It relies on consideration of a locally finite atlas. □

Remark 2.6 *If (M,e) has finite volume, then $L^p \subset L^q$ for all $p \leq q$. Indeed by the Hölder inequality it holds that*

$$||u||_{L^p} \leq Vol(M,e)^{\frac{1}{q'}}||u||_{L^q}, \quad \frac{1}{q} + \frac{1}{q'} = \frac{1}{p}. \tag{2.10}$$

[3] See for instance Friedman, A. (1976) *Partial Differential Equations*, Part 1 Section 10, Robert Krieger Publishing Company.

[4] Note that the inequality of the embedding theorem, $q \leq \frac{np}{n-(m-j)p}$ implies $a < 1$ if $j < m$. If $j = m$ then $a = 1$, a trivial result.

3 Weighted Sobolev spaces

3.1 *Definitions*

On non-compact manifolds, or manifolds with boundaries, one introduces variants of the previously defined Sobolev spaces by multiplying the pointwise norms by positive functions, called weights, to capture the behaviour "at infinity", or near the boundary. We will consider here the asymptotically Euclidean case only, which is of fundamental importance for the study of isolated systems.

The **Euclidean space** E^n is the manifold R^n endowed with the Euclidean metric which reads in canonical coordinates $\sum(dx^i)^2$. A C^∞, n-dimensional, Riemannian manifold (M, e) is called **Euclidean at infinity** if there exists a compact subset S of M such that $M - S$ is the disjoint union of a finite number of open sets U_i, and each (U_i, e) is isometric to the exterior of a ball in E^n. A manifold Euclidean at infinity is complete[5]. Each open set $U_i \subset M$ is sometimes called an **end** of M. If M is diffeomorphic to R^n it has only one end and we can then take for e the Euclidean metric. We denote by d the distance in the metric e from a point of M to a fixed point[6]. If (M, e) is a Euclidean space, one can choose $d = r$, the Euclidean distance to the origin.

Definition 3.1 *A* ***weighted Sobolev space*** $W^p_{s,\delta}$, $1 \leq p \leq \infty, s$ *a non-negative integer,* $\delta \in R$*, of tensors of some given type on the manifold* (M, e) *Euclidean at infinity is the space of functions, or tensor fields of some given type, on* (M, e) *with generalized derivatives of order* $\leq s$ *in the metric* e *such that* $\partial^m u(1 + d^2)^{\frac{1}{2}(\delta+m)} \in L^p$*, for* $0 \leq m \leq s$*. It is a Banach space with norm*

$$\|u\|_{W^p_{s,\delta}} = \left\{ \sum_{0 \leq m \leq s} \int_V | \partial^m u |^p \, (1 + d^2)^{\frac{1}{2}p(\delta+m)} d\mu_e \right\}^{1/p} . \tag{3.1}$$

The space $\mathcal{D}$ (C^∞ tensors with compact support in M) of tensors of the given type is dense in $W^p_{s,\delta}$ for any $1 \leq p < \infty$, any s and δ.

Definition 3.2

1. *A weighted* C^m_β *space is a Banach space of* m*-times continuously differentiable functions with norm*

$$\|u\|_{C^m_\beta} \equiv \sum_{0 \leq k \leq m} \sup_M (|\partial^k u|(1 + d^2)^{\frac{1}{2}(\beta+k)}). \tag{3.2}$$

2. *A weighted Hölder space* $C^{0,\alpha}_\beta$ *space is a Banach space with norm*

$$\|u\|_{C^{0,\alpha}_\beta} \equiv \|u\|_{C^0_\beta} + \sup_{(x,y) \in M \times M, d(x,y) < D} \frac{(|u(x) - u(y)|(1 + d^2)^{\frac{1}{2}\beta}}{d(x, y)}, \tag{3.3}$$

where D *is the injectivity radius of* M*.*

[5] Unless otherwise specified,a manifold is without boundary.
[6] Different choices of this point give equivalent norms.

3.2 *Embedding and multiplication properties*

It obviously holds that $W^p_{s,\delta} \subset W^p_{s',\delta'}$ if $s \geq s',\ \delta \geq \delta'$. It can be proved[7] that the imbedding of the space $W^p_{s,\delta}$ into $W^p_{s',\delta'}$ is compact if $s > s',\ \delta > \delta'$.

A straightforward application of the Hölder inequality and the integrability of $(1+|x|^2)^{-a/2}$ on R^n when $a > n$ gives the following theorem.

Theorem 3.3

1. *Embeddings.*
   (a) *If* $1 \leq p_1 \leq p_2 \leq \infty$, $\delta_2 - \delta_1 > n(\frac{1}{p_1} - \frac{1}{p_2})$, *then* $W^{p_2}_{0,\delta_2} \subset W^{p_1}_{0,\delta_1}$.
   (b) $C^0_\beta \subset W^p_{0,\delta}$ *if* $\beta > \delta + \frac{n}{p}$. (3.4)
2. *Continuous multiplication.*
   If $u \in W^{q_1}_{0,\delta_1}$, $v \in W^{q_2}_{0,\delta_2}$, $1 \leq q_1, q_2 \leq \infty$, $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{p} \geq 1$, *then* $uv \in W^p_{0,\delta}$ *and*:

$$||uv||_{W^p_{0,\delta}} \leq ||u||_{W^{q_1}_{0,\delta_1}} ||v||_{W^{q_2}_{0,\delta_2}}, \quad \delta_1 + \delta_2 = \delta. \tag{3.5}$$

More refined proofs[8] are required for the next theorem.

Theorem 3.4

1. *Sobolev embeddings.*
   (a) *If* $s > \frac{n}{p}$ *then*

$$W^p_{s+m,\delta} \subset C^{m,\alpha}_\beta \quad \textit{with} \quad \beta \leq \delta + \frac{n}{p}, \quad 0 < \alpha < s - \frac{n}{p}, \quad \alpha < 1. \tag{3.6}$$

   (b) *If* $s < \frac{n}{p}$, *then*

$$W^p_{s+m,\delta} \subset W^{np/(n-sp)}_{m,\delta+s}. \tag{3.7}$$

   (c) *If* $s = \frac{n}{p}$, *then*

$$W^p_{s+m,\delta} \subset W^q_{m,\delta+s} \quad \textit{for all } q < \infty. \tag{3.8}$$

2. *Continuous Sobolev multiplication:*

$$W^p_{s_1,\delta_1} \times W^p_{s_2,\delta_2} \subset W^p_{s,\delta} \quad \textit{if} \quad s \leq s_1, s_2,\ s < s_1 + s_2 - \frac{n}{p}, \quad \delta < \delta_1 + \delta_2 + \frac{n}{p}. \tag{3.9}$$

[7] Bancel, D. and Lacaze, J. (1976) *C. R. Acad. Sci. Paris*, **283**, 405–7.

[8] Choquet-Bruhat, Y. and Christodoulou, D. (1981) *Acta Mathematica*, **146**, 146–50 (see CB-DM2 p. 393). Bartnik, R. (1986) *Commun. Pures Appl. Math.*, **39**, 661–93, uses a different notation for the weight; he improves the embedding (a) from the strict inequality $\beta < \delta + n/p$ to the equality by his scaling method (see later our proof of the interpolation theorem).

By using the Bartnik scaling idea, we will prove the following theorem, generalizing a theorem given in Choquet-Bruhat and Christodoulou (1981) when $p = q$, $j = 1$, $m = 2$.

Theorem 3.5 *(Interpolation inequality) For any $\varepsilon > 0$, there is a number C_ε such that, for all $u \in W^p_{m,\delta}$, $1 \leq p \leq q \leq \infty$, $j < m$:*

$$||\partial^j u||_{W^q_{0,\delta+\frac{n}{p}-\frac{n}{q}+j}} \leq C\{\varepsilon||u||_{W^p_{m,\delta}} + C_\varepsilon||u||_{W^p_{0,\delta}}\}. \tag{3.10}$$

Proof We will use the Gagliardo–Nirenberg interpolation theorem (2.4). On a manifold Euclidean outside of a ball of radius $R_0 > 0$ and centre x_0 we denote by A_R the annulus $RR_0 < |x| < 2R_0R$, with $|x|$ the Euclidean distance from the point x to the origin of R^n and $R \geq 1$. We denote by $||.||_{W'^q_{0,\delta_q}(A_R)}$ and $||.||_{W'^q_{s,\delta_q}(A_R)}$ the norms on A_R defined by

$$||u||^q_{W'^q_{0,\delta_q}(A_R)} = \int_{|x|=R_0R}^{|x|=2R_0R} |x|^{q\delta_q}|u(x)|^q d^n x \tag{3.11}$$

and

$$||u||^q_{W'^q_{s,\delta_q}(A_R)} = \sum_{k=0}^{s} ||\partial^k u||^q_{W'^q_{s-k,\delta_q+k}(A_R)}. \tag{3.12}$$

We have, with $d(x)$ the distance between x and x_0 in the metric e,

$$||u||^q_{W^q_{0,\delta_q}(A_R)} = \int_{|x|=R_0R}^{|x|=2R_0R} |1 + d(x)^2|^{\frac{1}{2}q\delta_q}|u(x)|^q d^n x. \tag{3.13}$$

We define a function u_R by setting

$$u_R(y) = u(x), \quad x = Ry, \tag{3.14}$$

and we find by elementary calculus

$$||\partial^k u||^q_{W'^q_{0,\delta+k}(A_R)} = \int_{|y|=1}^{|y|=2} |y|^{q(\delta+k)} R^{q\delta+n}|\partial^k_y u_R(y)|^q d^n y; \tag{3.15}$$

that is:

$$||\partial^k u||_{W'^q_{0,\delta+k}(A_R)} = R^{\delta+\frac{n}{q}}||\partial^k u_R||_{W'^q_{0,\delta+k}(A_1)}. \tag{3.16}$$

Hence also

$$||u||_{W'^q_{m,\delta}(A_R)} \equiv \left\{ \sum_{0\leq k\leq m} ||\partial^k u||^q_{W'^q_{0,\delta+k}(A_R)} \right\}^{1/q} = R^{\delta+\frac{n}{q}}||u_R||_{W'^q_{m,\delta}(A_1)} \tag{3.17}$$

On the annulus A_1, the norms $W'^p_{s,\delta}$ and $W^p_{s,\delta}$ and W^p_s are equivalent for any s, p and δ. Using the interpolation inequality for u_R on A_1 gives therefore an inequality where C depends only on A_1, j, m, p, q

$$||\partial^j u_R||_{W^q_{0,\delta_q+j}(A_1)} \leq C||u_R||^a_{W^p_{m,\delta_p}(A_1)}||u_R||^{1-a}_{W^r_{0,\delta_r}(A_1)}. \tag{3.18}$$

Equation (3.16) gives then

$$R^{-(\delta_q+\frac{n}{q})}||\partial^j u||_{W^q_{0,\delta_q+j}(A_R)} \leq CR^{-a(\delta_p+\frac{n}{p})+(a-1)(\delta_r+\frac{n}{r})}||u||^a_{W^p_{m,\delta_p}(A_R)}||u||^{1-a}_{W^r_{0,\delta_r}(A_R)}. \tag{3.19}$$

We deduce from this inequality that, with C independent of R:

$$||\partial^j u||_{W^q_{0,\delta_q+j}(A_R)} \leq C||u||^a_{W^p_{m,\delta_p}(A_R)}||u||^{1-a}_{W^r_{0,\delta_r}(A_R)} \tag{3.20}$$

if

$$a\left(\delta_p + \frac{n}{p} - \delta_r - \frac{n}{r}\right) + \delta_r + \frac{n}{r} - \delta_q - \frac{n}{q} = 0 \tag{3.21}$$

Considering for simplicity the case of interest to us, $r = p$, $\delta_r = \delta_p = \delta$ gives $\delta_q = \delta + \frac{n}{p} - \frac{n}{q}$; that is

$$||\partial^j u||_{W^q_{0,\delta+\frac{n}{p}-\frac{n}{q}+j}(A_R)} \leq C||u||^a_{W^p_{m,\delta}(A_R)}||u||^{1-a}_{W^p_{0,\delta}(A_R)}. \tag{3.22}$$

Denote by A_0 the ball B_0. The norms $W^p_{s,\delta}(A_0)$ and $W^p_s(A_0)$ being equivalent, an inequality of the same type is valid in A_0.

The inequality (3.22) implies for norms on M that

$$\begin{aligned} ||\partial^j u||_{W^q_{0,\delta+\frac{n}{p}-\frac{n}{q}+j}} &= \left\{ \sum_{R=0,1,\ldots} ||\partial^j u||^q_{W^q_{\delta+\frac{n}{p}-\frac{n}{q}+j}(A_R)} \right\}^{\frac{1}{q}} \\ &\leq C \left\{ \sum_{R=0,1,\ldots} ||u||^{aq}_{W^p_{m,\delta}(A_R)}||u||^{(1-a)q}_{W^p_{0,\delta}(A_R)} \right\}^{\frac{1}{q}}. \end{aligned}$$

Hence

$$||\partial^j u||_{W^q_{0,\delta+\frac{n}{p}-\frac{n}{q}+j}} \leq C||u||^a_{W^p_{m,\delta}(A_R)}||u||^{(1-a)}_{W^p_{0,\delta}(A_R)}, \tag{3.23}$$

because, for any sequences of positive numbers it holds that

$$\left\{\sum_{R=0}^{\infty} a_R^q\right\}^{\frac{1}{q}} \leq \left\{\sum_{R=0}^{\infty} a_R^p\right\}^{\frac{1}{p}} \quad \text{if} \quad p \leq q \tag{3.24}$$

and

$$\sum_{R=0}^{\infty} a_R b_R \leq \left(\sum_{R=0}^{\infty} a_R\right)\left(\sum_{R=0}^{\infty} b_R\right). \tag{3.25}$$

The inequality (3.23) shows that if $a < 1$, i.e. if $j < m$, there exists for any given ε a number C_ε such that

$$||\partial^j u||_{W^q_{0,\delta+\frac{n}{p}-\frac{n}{q}+j}} \leq C\{\varepsilon||u||_{W^p_{m,\delta}} + C_\varepsilon||u||_{W^p_{0,\delta}}\}. \tag{3.26}$$

□

Remark 3.6 *The proof applies as well to the open submanifold M/B_{kR_0} by replacing the sum from zero to ∞ by a sum from zero to k. The constant k does not depend on k. This remark applies also to the embedding and multiplication theorems.*

The **Poincaré inequality**,[9] bounds the weighted norm of a function on R^n by a bound of a weighted norm of its derivative.

Theorem 3.7 *(Poincaré inequality) If $n \geq 3$, $\delta > -\frac{n}{2}$ there exists a constant C, depending on n, p and δ such that for any $u \in W^p_{1,\delta}$, $\delta > -\frac{n}{p}$*

$$||u||_{W^p_{0,\delta}} \leq C||\partial u||_{W^p_{0,\delta+1}}. \tag{3.27}$$

[9] For the proof see Lacaze, J. (1982) *C. R. Acad. Sci. Paris* case $p = 2$; Bartnik, R. (1986) *Commun. Pures Appl. Math.*, **39**, 661–93, general case.

APPENDIX II

SECOND-ORDER ELLIPTIC SYSTEMS ON RIEMANNIAN MANIFOLDS

1 Linear elliptic systems

Let (M, e) be a smooth Riemannian manifold. Denote by ∂ the covariant derivative in the metric e. A linear differential operator of order m from sections u of a tensor bundle E over (M, e) into sections of another such bundle F reads

$$Lu \equiv \sum_{k=0}^{m} a_k \partial^k u, \tag{1.1}$$

with ∂^k the covariant derivative of order k and a_k a linear map from tensor fields to tensor fields, given also by tensor fields over M.

The principal symbol of the operator L at a point $x \in M$, for a covector ξ at x, is the linear map $\sigma(\xi)$ from E_x to F_x (finite dimensional vector spaces) determined by the contraction of a_m with $(\otimes\xi)^m$. The operator is said to be elliptic if for each $x \in M$ and $\xi \in T_x^* M$ its principal symbol is an isomorphism from E_x onto F_x for all $\xi \neq 0$.

The **ellipticity constant** of the operator L is the infimum of the norms of the linear map $\sigma(\xi)$. If a_m is continuous and bounded on (M, e) this ellipticity constant is a positive number given by (the norms $|\ |$ are taken in the metric e)

$$\underset{x\in M, \xi\in T_x^*M, X_x\in E_x}{Inf} \frac{|(a_m(x) \otimes^m \xi)X_x|}{|\xi|^m |X_x|}.$$

Example 1.1 *The Laplace operator Δ_γ in a Riemannian metric γ on M acting on scalar functions has principal symbol*[1] $\gamma^{ij}\xi_i\xi_j$. *One verifies that it is elliptic. Its ellipticity constant is*

$$C_\gamma = \underset{x\in M, \xi\in T_x^*M}{Inf} \frac{\gamma^{ij}\xi_i\xi_j}{e^{ij}\xi_i\xi_j}.$$

Example 1.2 *The conformal Laplace operator in a Riemannian metric γ on M acting from vector fields X into vector fields is, with D the covariant derivative in the metric γ,*

$$\Delta_{\gamma,conf} X := D_i(\mathcal{L}_{\gamma,conf} X)^{ij} \equiv D_i \left(D^i X^j + D^j X^i - \frac{2}{n}\gamma^{ij} D_k X^k \right). \tag{1.2}$$

[1] The covariant derivatives in γ and e have the same principal part.

Its principal symbol at x, with $\xi \in T_x^* M$, is the linear mapping from the space of vectors X into itself given by

$$\xi^i \xi_i X^j + \xi_i \xi^j X^i - \frac{2}{n} \xi^j \xi_k X^k = Y^j \ . \tag{1.3}$$

This linear mapping is an isomorphism if $n \geq 2$ and $\xi \neq 0$ because $Y^j = 0$ implies

$$(\xi^i \xi_i)(X^j X_j) + \left(1 - \frac{2}{n}\right)(\xi_i X^i)^2 > 0 \tag{1.4}$$

if $\xi \neq 0$ and $X \neq 0$. The conformal Killing operator is elliptic.

We suppose that the metric γ is uniformly equivalent to the metric e; i.e. γ has a positive ellipticity constant and is uniformly bounded with respect to e. The conformal Laplace operator has then a positive ellipticity constant, given up to the equivalence of γ and e by

$$\begin{aligned} &\underset{x \in M, \xi \in T_x^* M, X \in T_x M}{Inf} \frac{|Y|_\gamma}{|X|_\gamma} \\ &= \underset{x \in M, \xi \in T_x^* M, X \in T_x M}{Inf} \frac{\{|\xi|_\gamma^2 |X|_\gamma^2 + \left(1 - \frac{2}{n}\right)(3 - \frac{2}{n})(\xi.X)_\gamma^2\}^{\frac{1}{2}}}{|\xi|_\gamma |X|_\gamma} = 1. \end{aligned} \tag{1.5}$$

We recall the following theorem[2] relative to homogeneous elliptic operators $\tilde{L}$ on bounded open sets of R^n, with $\tilde{\partial}$ the usual partial derivative

$$\tilde{L}u \equiv a_m \tilde{\partial}^m u. \tag{1.6}$$

Theorem 1.3 *(Douglis and Nirenberg) Let Ω be a bounded open set of R^n, with the Euclidean metric, and let a_m be continuous and bounded on Ω. Then the following estimate holds, for any $q > 1$:*

$$\|u\|_{\tilde{W}_m^q} \leq C_{a_m} \{\|\tilde{L}u\|_{\tilde{L}^q} + ||u||_{\tilde{L}^q}\} \tag{1.7}$$

where $\tilde{W}_s^q$ denotes Sobolev spaces defined by the Euclidean metric on Ω. The number C_{a_m} depends only on Ω, the C^0 norm of a_m, its modulus of continuity and its ellipticity constant.

In this appendix we will treat second-order systems, the only ones of relevance in the solution of the constraints that we consider. Such systems read in local

[2] Douglis, A. and Nirenberg, L. (1955) *Commun. Pure Appl. Math.*, **8**, 503–38.

coordinates x^i for the components u^A of a tensor (or spinor) field

$$(Lu)^B := a^{ij,B}_{2,A}\partial^2_{ij}u^A + a^{i,B}_{1,A}\partial_i u^A + a^B_{0,A}u^A = f^B. \tag{1.8}$$

2 Linear elliptic systems on compact M

2.1 *General second-order systems*

A compact manifold M can be covered by a finite number of charts with domains being bounded open sets of R^n. We endow it with a smooth Riemannian metric e uniformly equivalent in each chart to the Euclidean metric. We will sketch the proof of the following theorem.

Theorem 2.1 *Let (M, e) be a smooth compact orientable Riemannian manifold. Let*

$$Lu \equiv \sum_{k=0}^{2} a_k \partial^k u \tag{2.1}$$

be a second-order elliptic operator on (M, e). Suppose the coefficients of L are such that

$$a_2 \in W^p_2, \quad a_1 \in W^p_1, \quad a_0 \in L^p, \quad p > \frac{n}{2}. \tag{2.2}$$

Then:

1. *The operator L is a continuous mapping $W^q_2 \to L^q$ for any $1 < q \le p$.*
2. *The following estimate holds:*

$$\|u\|_{W^q_2} \le C_L\{\|Lu\|_{L^q} + ||u||_{W^q_1}\}. \tag{2.3}$$

The number C_L depends only on the norms of the $a_k^{'}s$ in their respective spaces and the ellipticity constant of a_2.

Proof

1. By the Sobolev embedding theorem we know that a_2 belongs to a Hölder space $C^{0,\alpha}$, since $W^p_2 \subset C^{0,\alpha}$, $0 < \alpha \le 2 - \frac{n}{p}$, if $p > \frac{n}{2}$. Therefore $a_2\partial^2 u \in L^q$ if $a_2 \in W^p_2$ and $u \in W^q_2$.

When $a_1 \in W^p_1$ and $\partial u \in W^q_1$ the Sobolev embedding theorem implies $a_1 \in L^{p_1}$ and $\partial u \in L^{q_1}$ if[3] $p < n$, $q < n$ and

$$p \le p_1 \le \frac{np}{n-p}, \quad q \le q_1 \le \frac{nq}{n-q}. \tag{2.4}$$

Then, by the Holder inequality

$$a_1\partial u \in L^r \quad \text{with} \quad \frac{1}{r} = \frac{1}{p_1} + \frac{1}{q_1} \ge \frac{n(p+q) - 2pq}{npq}. \tag{2.5}$$

[3] The property $a_1\partial u \in \mathrm{L}^q$ is trivial otherwise.

Therefore we have $a_1\partial u \in L^q$ if[4]

$$\frac{n(p+q)-2pq}{npq} \geq \frac{1}{q}.$$

This inequality is satisfied for any q if $p \geq \frac{n}{2}$.

Assume $a_0 \in L^p$ and $u \in W_2^q$; then $u \in L^{q_2}$, with $q_2 = +\infty$ if $q > \frac{n}{2}$, q_2 arbitrary large if $q_2 = \frac{n}{2}$, $\frac{1}{q_2} = \frac{1}{q} - \frac{2}{n}$ if $q < \frac{n}{2}$. In both cases $a_0 u \in L^r$, with $\frac{1}{r} = \frac{1}{p} + \frac{1}{q_2}$, hence $a_0 u \in L^q$ if $q \leq p$, $p > \frac{n}{2}$.

2. We use the Douglis–Nirenberg theorem on a bounded open set Ω of R^n.

We first treat the principal part $a_2\partial^2 u$ by considering a covering of M by a finite number of charts with bounded domains Ω_i, and using a partition of unity which consists of, functions $0 \leq \phi_i \leq 1$ with compact support in Ω_i such that $\sum_i \phi_i = 1$. We have $u = \sum u_i$, $u_i \equiv u\phi_i$. In any of the charts, e is uniformly equivalent to the Euclidean metric; we denote by $\tilde{\partial}$ the derivative in the Euclidean metric (i.e. the usual partial derivative), and we denote by $\tilde{W}_s^p$ the associated Sobolev norm. For fields with support in a chart the norms W_s^p and $\tilde{W}_s^p$ are equivalent; they are uniformly equivalent for the various charts, since they have compact domains and there are a finite number of them. The Douglis–Nirenberg theorem gives then the existence of a set of constants C_i, depending only on Ω_i and on a_2 through its $C^{0,\alpha}$ norm and the ellipticity constant of its representatives, such that

$$\|u_i\|_{W_2^q} \leq C_{i,a_2}\{\|\tilde{L}u_i\|_{L^q} + ||u||_{L^q}\}, \qquad \tilde{L}u_i \equiv a_2\tilde{\partial}^2 u_i. \tag{2.6}$$

The subadditivity of norms together with this inequality imply that, with C_{a_2} the maximum of the $C_i's$,

$$\|u\|_{W_2^q} \leq \sum_i ||u_i||_{W_2^q} \leq C_{a_2} \sum_i \{\|\tilde{L}u_i\|_{L^q} + ||u||_{L^q}\}. \tag{2.7}$$

On the other hand, we have an identity of the form

$$\tilde{L}u_i \equiv a_2\tilde{\partial}^2 u_i \equiv a_2\{\partial^2 u_i + S(i)\partial u_i\}, \tag{2.8}$$

where $S(i)$ are smooth coefficients linear in the connection of e.

Using the Leibniz rule we obtain that

$$\partial u_i = u\partial\phi_i + \partial u\phi_i, \qquad \partial^2 u_i = \partial^2 u\phi_i + 2\partial u\partial\phi_i + u\partial^2\phi_i. \tag{2.9}$$

Since $0 \leq \phi_i \leq 1$ we deduce from (2.9) that

$$|a_2\partial^2 u_i| \leq |a_2\partial^2 u| + |a_2(2\partial u\partial\phi_i + u\partial^2\phi_i)|. \tag{2.10}$$

Combining previous inequalities we find that there exists a number C, depending only on (M, e), such that

$$||\tilde{L}u_i||_{L^q} \leq C\{||a_2\partial^2 u||_{L^q} + ||a_2||_{C^0}||u||_{W_1^q}\}. \tag{2.11}$$

[4] Recall that on a compact manifold $L^q \subset L^r$ if $q \leq r$.

The inequalities (2.7), (2.8) and (2.10) imply therefore that

$$||u||_{W_2^q} \leq C_{a_2}\{||a_2\partial^2 u||_{L^q} + ||u||_{L^q}\}. \tag{2.12}$$

We then consider the general linear operator L, with coefficients satisfying the given hypotheses:

$$Lu \equiv a_2\partial^2 u + a_1\partial u + a_0 u. \tag{2.13}$$

Since $a_2 \in W_2^p$, with $p > \frac{n}{2}$, it belongs also to a Hölder space $C^{0,\alpha}$, $0 < \alpha \leq 2 - \frac{n}{p}$, and it holds that, with C a Sobolev constant of (M, e),

$$||a_2||_{C^{0,\alpha}} \leq C||a_2||_{W_2^p}. \tag{2.14}$$

The inequality (2.12) and the expression of Lu imply that

$$||u||_{W_2^q} \leq C_{a_2}\{||Lu||_{L^q} + ||a_1\partial u + a_0 u||_{L^q} + ||u||_{W_1^q}\}. \tag{2.15}$$

We estimate $a_1\partial u$ in L^q when $a_1 \in W_1^p$ and $u \in W_2^q$.

If $p > n$, then $a_1 \in C^0$ and $a_1\partial u \in L^q$ for all $q \leq n$, with

$$||a_1\partial u||_{L^q} \leq ||a_1||_{C^0}||\partial u||_{L^q}.$$

If $p < n$, we use the Hölder inequality

$$||a_1\partial u||_{L^q} \leq ||a_1||_{L^{p_1}}||\partial u||_{L^{q_1}}, \quad \frac{1}{p_1} + \frac{1}{q_1} = \frac{1}{q}; \tag{2.16}$$

where the Sobolev embedding theorem gives that

$$||a_1||_{L^{p_1}} \leq C||a_1||_{W_1^p}, \quad p_1 = \frac{np}{n-p}. \tag{2.17}$$

To take advantage of the property $u \in W_2^q$ without destroying the looked for inequality one uses the Gagliardo–Nirenberg interpolation inequality (see (I.2.7) in Appendix I). This inequality says that there exists a number C depending only on (M, e) such that

$$||\partial^j u||_{L^{q_1}} \leq C||u||^{\lambda}_{W_2^q}||u||^{1-\lambda}_{L^q} \tag{2.18}$$

if there exists a number λ such that

$$0 < \lambda := \frac{n}{2}\left\{\frac{j}{n} + \frac{1}{q} - \frac{1}{q_1}\right\} < 1. \tag{2.19}$$

In the case $j = 1$ and using the value of q_1 given by (2.16), this inequality reduces to

$$0 < 1 + \frac{n}{p_1} \equiv \frac{n}{p} < 2,$$

it is always satisfied if $p > \frac{n}{2}$.

We estimate $a_0 u$ in L^q

$$||a_0 u||_{L^q} \leq ||a_0||_{L^p}||u||_{L^{q_2}}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{q_2}. \tag{2.20}$$

By the interpolation inequality (2.18) in the case $j = 0$, q_1 replaced by q_2

$$||u||_{L^{q_2}} \leq C||\partial^2 u||^{\lambda}_{L^q}||u||^{1-\lambda}_{L^q}, \tag{2.21}$$

with

$$0 < \lambda := \frac{n}{2}\left\{\frac{1}{q} - \frac{1}{q_2}\right\} = \frac{n}{2p} < 1, \tag{2.22}$$

since $p > \frac{n}{2}$.

For any $\varepsilon > 0$, by elementary calculus, it holds that

$$||\partial^2 u||^{\lambda}_{L^q}||u||^{1-\lambda}_{L^q} \leq \lambda\varepsilon||\partial^2 u||_{L^q} + \frac{1}{(1-\lambda)\varepsilon^{\lambda/1-\lambda}}||u||_{L^q}; \tag{2.23}$$

hence there exists a number C such that

$$||a_1\partial u + a_0 u||_{L^q} \leq C\{||a_1||_{W_1^p} + ||a_0||_{L^p}\}\{\varepsilon||u||_{W_2^q} + C_\varepsilon||u||_{L^q}\} \tag{2.24}$$

Using this inequality together with (2.15) we see that we can choose $\varepsilon > 0$ small enough, depending on the bound of $||a_1||_{W_1^p} + ||a_0||_{L^p}$ and C_{a_2} so that the inequality

$$\|u\|_{W_2^q} \leq C_L\{\|Lu\|_{L^q} + ||u||_{W_1^q}\} \tag{2.25}$$

is satisfied. □

Note that when $n \leq 3$ then $2 > \frac{n}{2}$ and therefore $p = 2$ is in the range of the theorem.

The following corollary is important for working with evolution problems.

Corollary 2.2 *If the elliptic operator (2.1) is such that*[5]

$$a_2 \in H_s, \quad a_1 \in H_{s-1}, \quad a_0 \in H_{s-2}, \quad s > \frac{n}{2}, \quad s \geq 2,$$

then L is a continuous mapping $H_\sigma \to H_{\sigma-2}$ and for all $u \in H_\sigma$, $2 \leq \sigma \leq s$, it holds that

$$\|u\|_{H_\sigma} \leq C_L\{\|Lu\|_{H_{\sigma-2}} + ||u||_{H_{\sigma-1}}\}. \tag{2.26}$$

with C_L a number depending only on the norms of the $a_k^{'}s$ and the ellipticity constant of a_2.

Proof The mapping property comes immediately from the Sobolev multiplication theorem.

To prove the estimate we first remark that if a function f is in H_s, $s > \frac{n}{2}$, the Sobolev embedding theorem says that there exists $p > \frac{n}{2}$ such that $f \in W_2^p$.

[5] Maxwell, D. (2005) *J. Hyp. Diff. Eq.*, **2**, 521–46, considers non-integral s, in particular the case $n > \frac{3}{2}$ when $n = 3$. If we consider only integer s, then $s > \frac{n}{2}$ implies $s \geq 2$ as soon as $n \geq 3$.

The theorem implies therefore, under the new hypotheses on the coefficients of L, that

$$\|u\|_{H_2} \leq C_L\{\|Lu\|_{L^2} + ||u||_{H_1}\} \tag{2.27}$$

The inequality (2.26) can be proved inductively by differentiating the equation $s-2$ times. We think it will be more illuminating for the reader if we give the first two steps instead of a general formula.

1. We suppose $s \geq 3$. We denote by u_1 the first derivative[6] ∂u. The differentiated equation reads:

$$\partial Lu \equiv L_1 u_1 + \partial a_0 u, \tag{2.28}$$

where

$$L_1 u_1 := Lu_1 + (\partial a_2 + a_1)\partial u_1$$

is an operator whose coefficients have the same properties as those of Lu. The theorem gives

$$\|u_1\|_{H_2} \leq C_L\{\|\partial Lu\|_{L^2} + ||\partial a_0 u||_{L^2} + ||u_1||_{H_1}\}.$$

We estimate

$$||\partial a_0 u||_{L^2} \leq ||a_0||_{L^{p_1}} ||u||_{L^{p_2}}, \quad \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{2}, \tag{2.29}$$

using $a_0 \in H_{s-2}$, i.e. $\partial a_0 \in H_s - 3$.

If $s > 3 + \frac{n}{2}$, then $p_1 = \infty$ and (2.29) holds with $p_2 = 2$.

If $s < 3 + \frac{n}{2}$, then ∂a_0 is bounded in L^{p_1} with

$$\frac{1}{p_1} = \frac{1}{2} - \frac{s}{n} + \frac{3}{n},$$

and the inequality (2.29) holds with

$$\frac{1}{p_2} = \frac{s}{n} - \frac{3}{n}.$$

We apply the Gagliardo–Nirenberg interpolation inequality (I.2.8) with $j = 0$, $q = p_2$, $p = r = 2$, $m = 3$; we find that

$$||u||_{L^{p_2}} \leq C||u||^{\lambda}_{H_3}||u||^{1-\lambda}_{L^2} \leq \varepsilon||u||^{\lambda}_{H_3} + C_\varepsilon||u||^{!-\lambda}_{L^2}, \tag{2.30}$$

with

$$\lambda = \frac{n}{3}\left(\frac{1}{2} - \frac{1}{p_2}\right) = 1 + \frac{n}{3}\left(\frac{1}{2} - \frac{s}{n}\right). \tag{2.31}$$

[6] For simplicity in writing up the proof, we identify ∂ with $\tilde{\partial}$. Choosing e to be a non-Euclidean metric, as is necessary on non-flat manifolds, introduces complications in the writing but no conceptual change.

We check that $0 < \lambda < 1$ since we have supposed $s > \frac{n}{2}$ and $s < 3 + \frac{n}{2}$.

We leave the reader to check the bound of $\partial a_0 u$ in L^2 when $s = 3 + \frac{n}{2}$, where any finite $p_1 > 1$ is admissible.

2. We suppose $s \geq 4$. We differentiate (2.28) denoting $u_2 := \partial^2 u$. We find that

$$\partial^2 Lu \equiv L_2 u_2 + (\partial^2 a_1 + \partial a_0) u_1 + \partial^2 a_0 u, \tag{2.32}$$

where

$$L_2 u_2 := Lu_2 + (\partial a_2 + a_1)\partial u_2 + (\partial^2 a_2 + \partial a_1 + a_0) u_2.$$

The operator L_2 satisfies again the same properties as L and one estimates the supplementary term in (2.32) by using the interpolation inequality. Consider for instance the case $s < 4 + \frac{n}{2}$, and the term $\partial^2 a_0 u$. Since $\partial^2 a_0 \in H_{s-4}$, we have

$$\partial^2 a_0 \in L^{p_1}, \quad \frac{1}{p_1} = \frac{1}{2} - \frac{s}{n} + \frac{4}{n}$$

and

$$||\partial^2 a_0 u||_{L^2} \leq C_L ||u||_{L^{p_2}}, \quad \frac{1}{p_2} = \frac{s}{n} - \frac{4}{n}.$$

The interpolation inequality gives now

$$||u||_{L^{p_2}} \leq C ||u||_{H_4}^{\lambda} ||u||_{L^2}^{1-\lambda} \leq \varepsilon ||u||_{H_4}^{\lambda} + C_\varepsilon ||u||_{L^2}^{!-\lambda}, \tag{2.33}$$

with

$$\lambda = \frac{n}{4}\left(\frac{1}{2} - \frac{1}{p_2}\right) = 1 + \frac{n}{4}\left(\frac{1}{2} - \frac{s}{n}\right), \tag{2.34}$$

and, again, $0 < \lambda < 1$ since we have supposed $s > \frac{n}{2}$ and $s < 4 + \frac{n}{2}$.

The other cases and terms can be estimated similarly. The estimate thus found for $||u_2||_{H_2}$, together with the estimates for u_1 and u, give the inequality (2.26) for $\sigma = 4$ and $s > \frac{n}{2}$. The inequality for $\sigma \leq s$ can be proved along the same lines. □

Theorem 2.3 *Under the hypotheses of Theorem 2.1 it holds that*

1. *The operator L maps W_2^q into L^q with finite dimensional kernel and closed range.*
2. *If L is injective on W_2^q, then there is a number C_L such that for each u in W_2^q the following inequality holds:*

$$\|u\|_{W_2^q} \leq C_L \|Lu\|_{L^q}\,. \tag{2.35}$$

3. *If the formal adjoint ${}^*L$ of L satisfies the same hypothesis as L and is injective, then L is surjective from W_2^q onto L^q, and hence is an isomorphism if also injective.*

Corollary 2.4 *Under the hypothesis of Corollary 2.2, the theorem above holds with W_2^q and L^q replaced respectively by H_s and H_{s-2}, $s > \frac{n}{2}$.*

Proof

1. Recall that a Banach space is finite dimensional if the unit ball is a compact subset; i.e. if every sequence of elements of norm 1 is convergent. Consider a sequence $u_n \in W_2^q$, with $\|u_n\|_{W_2^q} \leq 1$. The unit ball of W_2^q is compact in the W_1^q topology. The sequence u_n admits therefore a subsequence, still denoted u_n, which converges in the W_1^q topology. If the sequence is in the kernel of the linear operator L, it satisfies by Theorem 2.1 the inequality

$$\|u_n - u_m\|_{W_2^q} \leq C_L ||u_n - u_m||_{W_1^q},$$

which shows that the subsequence converges also in the original topology of W_2^q. Hence the kernel is finite dimensional.

We denote by E a vector space complementary to the kernel $Ker(L)$ of L; the space E is closed since $Ker(L)$ is finite dimensional. Suppose there is no constant C such that

$$\|u\|_{W_2^q} \leq C\|Lu\|_{L^q} \quad \text{for all} \quad u \in E. \tag{2.36}$$

Then there exists a sequence $\{u_n\} \subset E$ such that $\|u_n\|_{W_2^q} = 1$ and $||Lu_n||_{W_1^q}$ tends to zero; hence Lu_n admits a Cauchy subsequence Lu_n in W_1^q. The subsequence u_n is then a Cauchy sequence in W_2^q by the above inequality; it converges to some function $\bar{u} \in E$, with $\|\bar{u}\|_{W_2^q} = 1$. By continuity $L\bar{u} = 0$. The existence of $\bar{u}$ contradicts the hypothesis that E is complementary to the kernel of L. The inequality (2.36) shows that the range of L as mapping from E, equivalently as mapping from W_2^q, is a closed subspace of L^q.

2. This is a particular case of (2.36) when the kernel of L reduces to zero.

3. The proof is classical; it uses as follows the duality relation. Let φ, ψ be arbitrary C^∞ functions on the compact manifold M. The following identity defines the formal adjoint L^*

$$\int_M L\psi\varphi\mu_e = \int_M \psi L^*\varphi\mu_e. \tag{2.37}$$

To show that the range of $L : W_2^q \to L^q$ is the entire space L^q it is sufficient to show that

$$\int_M L\psi\varphi\mu_e = 0 \quad \text{for all smooth } \psi \text{ implies } \varphi \equiv 0 \tag{2.38}$$

because the closure of $L\psi$ in L^q, that is the range of Lu, $u \in W_2^q$, is then identical to L^q, since L^q is the dual of a space $L^{q'}$, and C^∞ is dense in $L^{q'}$ since M is compact. The duality definition (2.37) shows that the condition (2.38) is equivalent to the the injectivity of L^*, which says that $L^*\varphi = 0$ implies $\varphi = 0$.

□

2.2 *Poisson operator*

The Poisson operator for a Riemannian metric γ on (M, e) acts on scalar functions u and reads, with S^h_{ij} the difference of the connections of e and γ

$$\triangle_\gamma u - au \equiv \gamma^{ij}(\partial^2_{ij}u + S^h_{ij}\partial_h u) - au. \tag{2.39}$$

We suppose that $\gamma \in W^p_2$, $p > \frac{n}{2}$; then $\gamma \in C^{0,\alpha}$. We denote by M^p_2 the open subset of symmetric 2-tensors in W^p_2 which are properly Riemannian metrics, namely such that they admit a positive ellipticity constant, $Inf_{\cup M_I} \det(\gamma_{ij}) > 0$, where γ_{ij} are the components of γ in a finite number of charts M_I covering M. The coefficient $a_2 \equiv \gamma^\#$, the contravariant tensor associated with γ, is then in $C^{0,\alpha}$ with a positive ellipticity constant. If $\gamma \in W^p_2$, then $S_{ij} \in W^p_1$, since linear in $\partial\gamma$. The coefficient $a_1 \equiv \gamma^{ij}S_{ij}$ is therefore in W^p_1 if $p > \frac{n}{2}$. The hypotheses of Theorem 2.1 are satisfied if also $a \in L^p$. Hence:

Lemma 2.5 *If $\gamma \in M^p_2$ and $a \in L^p$, $p > \frac{n}{2}$, the Poisson operator is a continuous mapping $W^q_2 \to L^q$ so long as $1 < q \leq p$.*

We now prove a uniqueness lemma[7].

Lemma 2.6 *Let $\gamma \in M^p_2$, $p > \frac{n}{2}$, let $q \geq \frac{2n}{n+2}$, and let $a \in L^p$ with $a \geq 0$, $a \not\equiv 0$ (i.e. $a > 0$ on a subset of M of positive measure). Then the Poisson operator $\triangle_\gamma - a$ is injective on W^q_2.*

Proof The proof for C^2 functions results from the maximum principle if the coefficients are bounded. In the more general case one proceeds as follows: on a compact manifold, if[8] u_n is a C^2 function and $\gamma_n \in C^1$, the following identity is obtained by a straightforward integration by parts:

$$\int_M u_n(\Delta_{\gamma_n} u_n)\mu_{\gamma_n} \equiv -\int_M (\gamma^{ij}_n \partial_i u_n \partial_j u_n)\mu_{\gamma_n}. \tag{2.40}$$

This identity applied to a C^2 solution of $\Delta_{\gamma_n} u_n - a u_n = 0$, implies that $\partial_i u_n \equiv 0$ on M, and $u_n = 0$ on the subset on which $a > 0$. Hence u_n =constant, and $u_n \equiv 0$.

Suppose now that $\gamma \in M^p_2$, $p > \frac{n}{2}$, and $u \in W^p_2$. Since C^2 is dense in W^p_2 we can approximate γ and a by smooth sequences γ_n and a_n converging to γ and a, respectively in M^p_2 and L^p. We approximate $u \in W^q_2$ by a smooth sequence u_n. We already know by Lemma 2.5 that $\Delta_{\gamma_n} u_n$ converges in L^q to $\Delta_\gamma u$. Since the measures μ_{γ_n} and μ_γ are equivalent to μ_e, we will have shown that the integrals on both sides of (2.40) converge to limits if there are continuous

[7] The condition $p > n/2$ for u is not necessary. A sufficient condition is $p \geq \frac{2n}{n+2}$ if $n > 2$, $p > 1$ if $n = 2$. We limit ourselves to $p > \frac{n}{2}$ because in studying non-linear equations we will need u to be continuous.

[8] We use n both for the dimension of space and as an index labelling elements of a sequence. The alphabet has not enough elements to avoid this double meaning, which should not create confusion.

embedding $W_2^q \times L^q \subset L^1$ and $W_1^q \times W_1^q \subset L^1$. If $q > \frac{n}{2}$, then $W_2^q \subset L^\infty$, hence $W_2^q \times L^q \subset L^q$, but $L^q \subset L^1$ since the manifold M is compact. If $q < \frac{n}{2}$, then $W_2^q \subset L^{q_2}$, $q_2 = \frac{nq}{n-2q}$; hence $W_2^q \times L^q \subset L^r$ for $\frac{1}{r} = \frac{1}{q} + \frac{1}{q_2} = 2(\frac{1}{q} - \frac{1}{n})$. We have $r \geq 1$ if $\frac{1}{q} - \frac{1}{n} \leq \frac{1}{2}$. Similar use of Sobolev and Hölder inequalities give the result $W_1^q \times W_1^q \subset L^r$, with $r \geq 1$ if $q \geq \frac{2n}{n+2}$. The convergence in L^1 of $a_n u_n^2$ to au^2, obvious when $q > \frac{n}{2}$, results when $\frac{2n}{n+2} \leq q < \frac{n}{2}$, from the inequality $\frac{1}{p} + \frac{2}{q_2} < 1$ if $p > \frac{n}{2}$. We have thus proved that the equality

$$\int_M u(\Delta_\gamma u - au)\mu_\gamma \equiv -\int_M (\gamma^{ij}\partial_i u \partial_j u + au^2)\mu_\gamma \tag{2.41}$$

holds under the given conditions on γ, a and u. It implies $u \equiv 0$ if $\Delta_\gamma u - au = 0$ and $a \geq 0, a \not\equiv 0$. □

Theorem 2.7 *The Poisson operator $\triangle_\gamma - a$ on scalar functions for a metric γ on a smooth compact Riemannian manifold (M, e), with $\gamma \in M_2^p$, $a \in L^p$ and $p > \frac{n}{2}$, is an isomorphism from W_2^q onto L^q so long as $a \geq 0$, $a \not\equiv 0$, $\frac{2n}{n+2} \leq q \leq p$, and $n > 2$. If $n = 2$ the lower bound of q is the strict inequality.*

Corollary 2.8 *The Poisson operator is an isomorphism $H_s \to H_{s-2}$ if $\gamma \in M_s$, $a \in H_{s-2}$, $s > \frac{n}{2}$, $a \geq 0$ and $a \not\equiv 0$.*

Proof Under the given hypotheses the Poisson operator is self-adjoint relative to the metric γ: two integrations by parts for smooth enough γ, u and v, and taking limits, show that under the given hypotheses, as in the proof of Lemma 2.6, the following identity holds for $u, v \in W_2^q$

$$\int_M v(\Delta_\gamma u - au)\mu_\gamma \equiv \int_M u(\Delta_\gamma v - av)\mu_\gamma. \tag{2.42}$$

The isomorphism Theorem 2.3 and its Corollary[9] 2.4 apply therefore to our Poisson operator. □

2.3 *Conformal Laplace operator*

The conformal Laplace operator for a Riemannian metric γ is the divergence of the conformal Lie derivative acting on vector fields; that is, it is given by, with D the covariant derivative for the metric γ,

$$(\Delta_{\gamma,conf} X)^j := D_i \left(D^i X^j + D^j X^i - \frac{2}{n}\gamma^{ij} D_k X^k \right). \tag{2.43}$$

In solving the momentum constraint we will use the following theorem.

Theorem 2.9 *Suppose that $\gamma \in M_2^p$, $p > \frac{n}{2}$, $q > 1$, $\frac{2n}{n+2} \leq q \leq p$. Then:*

1. *The kernel of the conformal Laplace operator $\Delta_{\gamma,conf}$ in the space of W_2^q vector fields is the space of W_2^q conformal Killing vector fields.*

[9] Note that $2 \geq \frac{2n}{n+2}$ for any n.

2. *If* (M, γ) *admits no* W_2^q *conformal Killing vector field, then*
   (a) $\Delta_{\gamma,conf}$ *is an isomorphism from* W_2^q *onto* L^q.
   (b) *If* $\gamma \in M_s$, $s > \frac{n}{2}$, *then* $\Delta_{\gamma,conf}$ *is an isomorphism from* H_s *onto* H_{s-2}.

Proof

1. On a compact manifold integration by parts shows that the following identity holds for smooth vector fields X:

$$\int_M X_j(\Delta_{\gamma,conf}X)^j\mu_\gamma \equiv \int_M X_j D_i\left(D^iX^j + D^jX^i - \frac{2}{n}\gamma^{ij}D_kX^k\right)\mu_\gamma \equiv$$
$$-\int_M \left(D^iX^j + D^jX^i - \frac{2}{n}\gamma^{ij}D_kX^k\right)\left(D_iX_j + D_jX_i - \frac{2}{n}\gamma_{ij}D_lX^l\right)\mu_\gamma. \quad (2.44)$$

Proofs analogous to the proofs given for the Laplace operator show that both sides of this identity are defined under the hypothesis made on γ and X, and are limits of the corresponding expressions constructed with smooth vector fields. Hence the identity holds and $\Delta_{\gamma,conf}X = 0$ implies, for W_2^p vector fields,

$$D_iX_j + D_jX_i - \frac{2}{n}\gamma_{ij}D_lX^l = 0. \quad (2.45)$$

2. This follows from Theorem 2.3 and its corollary, since $\Delta_{\gamma,conf}$ is self-adjoint.

□

3 Asymptotically Euclidean manifolds

3.1 *Definitions*

Manifolds (M, e) which are Euclidean at infinity and weighted Sobolev spaces $W^p_{s,\delta}$ on such manifolds have been defined in Appendix I. They are Banach spaces with norm

$$\|u\|_{W^p_{s,\delta}} = \left\{\sum_{0\le m\le s}\int_V |\,\partial^m u\,|^p\,(1+d^2)^{\frac{1}{2}p(\delta+m)}d\mu_e\right\}^{1/p}. \quad (3.1)$$

If (M, e) is a Euclidean space one can choose $d = r$, the Euclidean distance to the origin.

We recall the following theorem relative to homogeneous elliptic operators $\tilde{L}$ on R^n, with $\tilde{\partial}$ the usual partial derivative.

$$\tilde{L}u \equiv a_m\tilde{\partial}^m u. \quad (3.2)$$

Theorem 3.1 *(Nirenberg and Walker). Let* a_m *be defined on* R^n *and be such that*

$$a_m - A_m \in C^{0,\alpha}_\beta, \quad \beta > 0 \quad (3.3)$$

where $L_\infty \equiv A_m\tilde{\partial}^m$ is an elliptic operator with constant coefficients. Then the following estimate holds, for any $q > 1$ and any δ, if $m < n$

$$\|u\|_{\tilde{W}^q_{m,\delta}} \leq C_{a_m}\{\|\tilde{L}u\|_{\tilde{W}^q_{0,\delta+m}} + ||u||_{\tilde{W}^q_{0,\delta}}\} \tag{3.4}$$

where $\tilde{W}^q_{s,\delta}$ denotes weighted Sobolev spaces defined by the Euclidean metric on R^n. The number C_{a_m} depends only on A_m, the ellipticity constant of a_m and the $C^{0,\alpha}_\beta$ norm of $a_m - A_m$.

3.2 *Second-order linear elliptic systems*

We consider linear second-order differential elliptic systems on manifolds which are Euclidean at infinity and of dimension[10] $n \geq 3$.

$$Lu := a_2\partial^2 u + a_1\partial u + a_0 u = f. \tag{3.5}$$

We will obtain solutions with a minimum degree of strong differentiability[11], namely solutions which are in $W^p_{2,\delta}$, $p > \frac{n}{2}$, and also solutions in $H_{s,\delta}$, $s > \frac{n}{2}$, $s \geq 2$, which are important for a tie-in with the evolution problem[12]. The steps of the proofs will be the same as in the case of a compact manifold; new technical lemmas have to be proved.

Lemma 3.2 *Suppose that $a_1 \in W^p_{1,\delta+1}$, $a_0 \in W^p_{0,\delta+2}$, $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$ and that $u \in W^q_{2,\delta_q}$ for some δ_q and $1 < q \leq p$ Then for any $\varepsilon > 0$ there exists C_ε such that*

1. $$||a_1\partial u||_{W^q_{0,\delta_q+2}} \leq ||a_1||_{W^p_{1,\delta+1}}\{\varepsilon||u||_{W^q_{2,\delta_q}} + C_\varepsilon||u||_{W^q_{1,\delta_q}}\} \tag{3.6}$$
2. $$||a_0 u||_{W^q_{0,\delta q+2}} \leq ||a_0||_{W^p_{0,\delta+2}}\{\varepsilon||u||_{W^q_{2,\delta_q}} + C_\varepsilon||u||_{W^q_{1,\delta_q}}\}. \tag{3.7}$$

Proof

1. The Hölder inequality (I.3.5) gives

$$||a_1\partial u||_{W^q_{0,\delta_q+2}} \leq ||a_1||_{W^{p_1}_{0,\delta_{p_1}}}||\partial u||_{W^{q_1}_{0,\delta_{q_1}}} \tag{3.8}$$

with

$$\frac{1}{p_1} + \frac{1}{q_1} = \frac{1}{q}, \quad \delta_{p_1} + \delta_{q_1} = \delta_q + 2. \tag{3.9}$$

[10] The Nirenberg–Walker theorem applies to such systems only if $n > 2$ (see Note before Lemma 3.5).

[11] Choquet-Bruhat, Y., Isenberg, J., and York, J. (2000) *Phys. Rev. D*, **3**(8), 83–105, case $p > \frac{n}{2}$, $s \geq 2$.

[12] Choquet-Bruhat, Y. and Christodoulou, D. (1981) *Acta Mathematica*, **146**, 129–50 treat the case $H_{s,\delta}$ with integer $s > \frac{n}{2}$.
Maxwell, D. (2006) *J. Reine Angew. Math.*, 1–29 extends the results to $H_{s,\delta}$, $s > \frac{n}{2}$ not necessarily an integer, by using refined methods.

(a) Suppose that $p > n$. Then $a_1 \in W^p_{1,\delta+1} \subset C^0_{\beta+1}$ for all $\beta \leq \delta + \frac{n}{p}$, hence for some $\beta > 0$. An elementary estimate of the integral gives

$$||a_1 \partial u||_{W^q_{0,\delta_q+2}} \leq ||a_1||_{C^0_1} ||\partial u||_{W^q_{0,\delta_q+1}} \leq ||a_1||_{C^0_{\beta+1}} ||u||_{W^q_{1,\delta_q}}. \tag{3.10}$$

(b) Suppose $\frac{n}{2} < p < n$. We have (I.3.7)

$$a_1 \in W^p_{1,\delta+1} \subset W^{p_1}_{0,\delta_{p_1}} \quad \text{with} \quad p_1 = \frac{np}{n-p}, \quad \delta_{p_1} = \delta + 2. \tag{3.11}$$

Hence

$$||a_1 \partial u||_{W^q_{0,\delta_q+2}} \leq C||a_1||_{W^p_{1,\delta+1}} ||\partial u||_{W^{q_1}_{0,\delta_{q_1}}}. \tag{3.12}$$

with, by (3.9) and (3.11)

$$\frac{1}{q_1} = \frac{1}{q} - \frac{1}{p} + \frac{1}{n}, \quad \delta_{q_1} = \delta_q - \delta. \tag{3.13}$$

We apply the interpolation inequality (I.3.10). We find

$$||\partial u||_{W^{q_1}_{0,\delta_1}} \leq C\{\varepsilon ||u||_{W^q_{2,\delta_q}} + C_\varepsilon ||u||_{W^q_{0,\delta_q}}\} \tag{3.14}$$

with

$$\delta_1 = \delta_q + \frac{n}{q} - \frac{n}{q_1} + 1 = \delta_{q_1} + \delta + \frac{n}{p}. \tag{3.15}$$

Since by hypothesis $\delta + \frac{n}{p} > 0$, we have

$$\delta_1 > \delta_{q_1}; \tag{3.16}$$

therefore (3.14) implies the inequality

$$||\partial u||_{W^{q_1}_{0,\delta_{q_1}}} \leq C\{\varepsilon ||u||_{W^q_{2,\delta_q}} + C_\varepsilon ||u||_{W^q_{0,\delta_q}}\}, \tag{3.17}$$

which completes the proof.

2. Multiplication and interpolation theorems permit us again to complete the proof. Indeed we have (multiplication)

$$||a_0 u||_{W^q_{0,\delta_q+2}} \leq ||a_0||_{W^p_{0,\delta+2}} ||u||_{W^{q_0}_{0,\delta_{q_0}}}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{q_0}, \ \delta_q = \delta + \delta_{q_0}$$

and (interpolation)

$$||u||_{W^{q_0}_{0,\delta_0}} \leq C\{\varepsilon ||u||_{W^q_{2,\delta_q}} + C_\varepsilon ||u||_{W^q_{0,\delta_q}}\}, \quad \delta_0 = \delta_q + \frac{n}{q} - \frac{n}{q_0}.$$

Using the above equalities, we see that

$$\delta_0 = \delta_q + \frac{n}{p} = \delta_{q_0} + \delta + \frac{n}{p} > \delta_{q_0}. \tag{3.18}$$

which completes the proof. □

The following theorem corresponds to Theorem 2.1 of the compact case.

Theorem 3.3 *Let (M, e) be a smooth Riemannian manifold which is Euclidean at infinity. Let*

$$Lu \equiv a_2 \partial^2 u + a_1 \partial u + a_0 u \tag{3.19}$$

be a second-order elliptic operator on (M, e). Suppose the coefficients of L satisfy the following hypotheses with $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$:

$$a_2 - A \in W^p_{2,\delta}, \quad a_1 \in W^p_{1,\delta+1}, \quad a_0 \in W^p_{0,\delta+2}. \tag{3.20}$$

where $A\partial^2$ is an elliptic operator on M with C^∞ coefficients, constant at each end of (M, e). Then, for any $1 < q \leq p$, and any δ_q

1. *The operator L is a continuous mapping from W^q_{2,δ_q} into W^q_{0,δ_q+2}.*
2. *The following inequality holds, with C_L a number depending only on A and on the norms of $a_2 - A$, a_1 and a_0*

$$\|u\|_{W^q_{2,\delta_q}} \leq C_L \{ \|Lu\|_{W^q_{0,\delta_q+2}} + ||u||_{W^q_{1,\delta_q}} \} . \tag{3.21}$$

Proof

1. Results from Lemma 3.2, after noticing that $a_2 - A \in W^p_{2,\delta}$, $p > \frac{n}{2}$, $\delta - \frac{n}{p}$ implies that a_2 is continuous and bounded on M.

2. We first consider a homogeneous operator $L \equiv a_2 \partial^2$. Since (M, e) is Euclidean at infinity, its ends Ω_I are contained in the complement of a geodesic ball B_R of finite radius R and centre some point x_0. We cover a larger ball, for instance B_{2R} by a finite number of open sets Ω_i, relatively compact smooth Riemannian manifolds with smooth boundaries. We denote by ϕ_i, ϕ_I a smooth partition of unity on M such that

$$\phi_i = 0 \quad \text{in} \quad M \backslash B_{2R}, \tag{3.22}$$

$$\phi_I = 1 \ \text{ in } \ \Omega_I \cap M \backslash B_{2R}, \quad \phi_I = 0 \ \text{ in } \ M \backslash B_R \quad \text{and in } \Omega_J \ \text{ if } \ I \neq J. \tag{3.23}$$

For every u we have

$$u \equiv \sum_i u_i + \sum_I u_I, \quad u_i =: \phi_i u, \quad u_I =: \phi_I u. \tag{3.24}$$

Each (Ω_i, e) is isometric with a bounded open set of R^n endowed with a metric uniformly equivalent to the Euclidean metric. We denote by $\tilde{\partial}$ the derivative in the Euclidean metric (i.e. the usual partial derivative), by $\tilde{W}^q_s$ the associated

Sobolev norm. For fields with support in Ω_i the norms $\tilde{W}^q_s$ and W^q_s are equivalent, also with the norm $W^q_{s,\delta}$ whatever δ is. These norms are uniformly equivalent for the various Ω_i, since there is a finite number of them. The Douglis–Nirenberg theorem then guarantees the existence of a constant C_i, depending only on Ω_i and on a_2 through the $C^{0,\alpha}$ norm and the ellipticity constant of its representative in Ω_i, such that

$$\|u_i\|_{W^q_2(\Omega_i)} \leq C_{i,a_2}\{\|\tilde{L}u_i\|_{L^q(\Omega_i)} + ||u_i||_{L^q(\Omega_i)}\}, \qquad \tilde{L}u_i \equiv a_2\tilde{\partial}^2 u_i. \tag{3.25}$$

The subadditivity of norms together with this inequality and $||u_i||_{L^q(\Omega_i)} \leq ||u||_{L^q(\Omega_i)}$, since $0 \leq \phi_i \leq 1$, imply that, with C_{a_2} the maximum of the C_{i,a_2}'s, and Ω the union of the $\Omega_i's$,

$$\left\|\sum_i u_i\right\|_{W^q_2(\Omega)} \leq \sum_i ||u_i||_{W^q_2(\Omega_i)} \leq C_{a_2}\left\{\sum_i \|\tilde{L}u_i\|_{L^q(\Omega_i)} + ||u||_{L^q(\Omega)}\right\}. \tag{3.26}$$

On the other hand, we have an identity of the following form, with $S(i)$ smooth coefficients linear in the connection of e:

$$\tilde{L}u_i := a_2\tilde{\partial}^2 u_i \equiv a_2\{\partial^2 u_i + S(i)\partial u_i\}. \tag{3.27}$$

Using the Leibnitz rule we obtain that

$$\partial u_i = u\partial\phi_i + \partial u\phi_i, \qquad \partial^2 u_i = \partial^2 u\phi_i + 2\partial u\partial\phi_i + u\partial^2\phi_i. \tag{3.28}$$

Since $0 \leq \phi_i \leq 1$ we deduce from these relations the following pointwise inequality in Ω_i:

$$|a_2\partial^2 u_i| \leq |a_2\partial^2 u| + |a_2(2\partial u\partial\phi_i + u\partial^2\phi_i)|. \tag{3.29}$$

Combining previous inequalities, we find that there exists a number C depending only on (M, e), such that

$$||\tilde{L}u_i||_{L^q(\Omega_i)} \leq C\{||a_2\partial^2 u||_{L^q(\Omega_i)} + ||a_2||_{C^0(\Omega_i)}||u||_{W^q_1(\Omega_i)}\}. \tag{3.30}$$

The inequality (3.25) implies then that

$$\left\|\sum_i u_i\right\|_{W^q_{2,\delta_q}} \leq C_{a_2}\{||a_2\partial^2 u||_{W^q_{0,\delta_q+2}} + ||u||_{W^q_{1,\delta_q}}\}. \tag{3.31}$$

We proceed similarly with $\sum_I u_I$, applying the Nirenberg–Walker inequality to u_I. We find that (in Ω_I, where u_I has its support, the metric e is identical with the Euclidean metric, i.e. $\partial = \tilde{\partial}$)

$$\|u_I||_{W^p_{2,\delta_q}} \leq C_{a_2}\{\|a_2\partial^2 u_I\|_{W^p_{0,\delta_q+2}} + ||u_I||_{W^p_{0,\delta_q}}\} \tag{3.32}$$

A pointwise inequality similar to (3.29), using now the properties of ϕ_I, in particular the fact that $\partial\phi_I = 0$ in $M\backslash B_{2R}$, shows that $\sum_I u_I$ satisfies an inequality

of the same form as $\sum_i u_i$, where C_{a_2} depends on (M, e), on the ellipticity constant of a_2 and on the $C^{0,\alpha}_\beta$ norm of $a_2 - A$. Since $a_2 - A \in W^p_{2,\delta}$, with $p > \frac{n}{2}$, it belongs also to a Hölder space $C^{0,\alpha}_\beta$, $0 < \alpha \leq 2 - \frac{n}{p}$, for all β such that $\beta \leq \delta + \frac{n}{p}$, and it holds that, with C a Sobolev constant of (M, e),

$$||a_2 - A||_{C^{0,\alpha}_\beta} \leq C||a_2 - A||_{W^p_{2,\delta}}. \tag{3.33}$$

The proof for the operator $a_2\partial^2$ is complete.

We now consider the general linear operator L, with coefficients satisfying the given hypotheses:

$$Lu \equiv a_2\partial^2 u + a_1\partial u + a_0 u. \tag{3.34}$$

The inequality proved for $a_2\partial^2$ implies that

$$||u||_{W^q_{2,\delta_q}} \leq C_{a_2}\{||Lu||_{W^q_{0,\delta_q+2}} + ||a_1\partial u + a_0 u||_{W^q_{0,\delta_q+2}} + ||u||_{W^p_{1,\delta_q}}\}. \tag{3.35}$$

The given inequality follows then from Lemma 3.2 by choosing, in this lemma, ε small enough. □

Corollary 3.4 *If the elliptic operator (3.18) is such that*[13]

$$a_2 - A \in H_{s,\delta}, \quad a_1 \in H_{s-1,\delta+1}, \quad a_0 \in H_{s-2,\delta+2}, \quad s > \frac{n}{2}, \quad s \geq 2, \quad \delta > -\frac{n}{2}$$

then L is a continuous mapping $H_{s,\delta} \to H_{s-2,\delta+2}$ and for all $u \in H_{s,\delta}$,

$$\|u\|_{H_{s,\delta}} \leq C_L\{\|Lu\|_{H_{s-2,\delta+2}} + ||u||_{H_{s-1,\delta+1}}\}. \tag{3.36}$$

with C_L a number depending only on A, the norms of the a_ks and the ellipticity constant of a_2.

Proof Recall the embedding (Appendix I, Theorem 3.4)

$$H_{s,\delta} \subset W^p_{2,\delta_p} \quad \text{if} \quad s < \frac{n}{2} + 2, \quad p = \frac{2n}{n - 2s + 4} \qquad \delta_p = \delta + s - 2.$$

If $s > \frac{n}{2}$ we have $p > \frac{n}{2}$. The condition on δ_p gives, if $\delta > -\frac{n}{2}$, $\delta_p > -\frac{n}{2} + s - 2$, that is, for the considered p, $\delta_p > -\frac{n}{p}$. The operator L satisfies therefore the hypothesis of the theorem. The proof of the corollary can be made along the same lines as for Corollary 2.2 (compact case), using multiplication, and interpolation in weighted Sobolev spaces. □

To prove an analog of Theorem 2.3 we need the following lemma[14].

[13] Maxwell, D. (2005) *J. Hyp. Diff. Eq.*, **2**, 521–46, considers non-integral s, in particular the case $s > \frac{3}{2}$ when $n = 3$. If we consider only integer s, then $s > \frac{n}{2}$ implies $s \geq 2$ so long as $n \geq 3$.

[14] In the asymptotically Euclidean case the dimension $n = 2$ is excluded from our results, due to the lack of fall off at infinity of $\log r$, an elementary solution of the Laplacian.

Lemma 3.5 *If $n \geq 3$, the operator $A\partial^2$ is an isomorphism $W^q_{2,\delta_q} \to W^q_{0,\delta_q+2}$ for each δ_q such that $-\frac{n}{q} < \delta_q < n-2-\frac{n}{q}$.*

Proof It follows the lines of the proofs of lemmas 5.1 and 5.2 of Choquet-Bruhat and Christoulou, which use the John's Green function of operators on R^n. □

Lemma 3.6

1. *Suppose that $u \in W^q_{2,\delta_q}$ is a solution of the equation*

$$Lu \equiv a_2\partial^2 u + a_1\partial u + a_0 u = f \tag{3.37}$$

where L satisfies the hypothesis of Theorem 3.3 and where $f \in W^q_{0,\tilde{\delta}+2}$, $\delta_q \leq \tilde{\delta}$. Then u is in fact in $W^q_{2,\tilde{\delta}}$, as long as $\tilde{\delta} < n-2-\frac{n}{q}$.

2. *If $u \in W^q_{2,\delta_q}$ is in the kernel of the operator L, then $u \in W^q_{2,\tilde{\delta}}$ for all $\tilde{\delta}$ such that $\tilde{\delta} < n-2-\frac{n}{q}$.*

Proof

1. The solution u satisfies the equation

$$A\partial^2 u = (A-a_2)\partial^2 u - a_1\partial u - a_0 u + f = 0. \tag{3.38}$$

If $u \in W^q_{2,\delta_q}$, then under the hypothesis made on L and on f, the right-hand side belongs to a space W^q_{0,δ_1+2} for some δ_1 such that $\tilde{\delta} \geq \delta_1 > \delta_q$. The isomorphism theorem for $A\partial^2$ shows that $u \in W^q_{2,\delta_1}$ if $\delta_1 < n-2-\frac{n}{q}$. An integration argument completes the proof.

2. If u is in the kernel of the operator, then $f \equiv 0 \in W^q_{0,\tilde{\delta}}$ whatever be $\tilde{\delta}$. □

The following theorem and corollary can be proved, using these results, by functional methods analogous to methods used in the compact case.

Theorem 3.7 *Under the hypothesis of Theorem 3.3 and if moreover*

$$-\frac{n}{q} < \delta_q < -2+n-\frac{n}{q} \tag{3.39}$$

1. *The operator L has finite dimensional kernel and closed range.*

2. *If L is injective on W^q_{2,δ_q}, then there is a number C_L such that for each u in W^q_2 the following inequality holds:*

$$\|u\|_{W^q_{2,\delta_q}} \leq C_L \|Lu\|_{W^q_{0,\delta+2}} . \tag{3.40}$$

3. *If the formal adjoint ${}^*L$ of L satisfies the same hypothesis as L and if L^* is injective, then L is surjective from $W^q_{2,\delta}$ onto $W^q_{0,\delta+2}$, and hence L is an isomorphism if it is also injective.*

Corollary 3.8 *Theorem 3.7 holds with* $H_{s,\delta}$, $H_{s-1,\delta+1}$, $H_{s-2,\delta+2}$, $s > \frac{n}{2}$, $-\frac{n}{2} < \delta < -2 + \frac{n}{2}$, *replacing* $W^p_{2,\delta}$, $W^p_{1,\delta+1}$, $W^p_{0,\delta+2}$ *in the hypothesis on the coefficients* $A - a_2$, a_1, a_0, *and also replacing* W^q_{2,δ_q}, W^q_{1,δ_q+1}, $W^q_{0,\delta q+2}$ *for the unknown* u.

4 Special systems

4.1 *Poisson operator*

We apply Theorem 3.7 to the Poisson operator $\triangle_\gamma - a$ on scalar functions, after proving a criterion for its injectivity.

We prove injectivity by integration of the product with u, but we first prove a lemma.

Lemma 4.1 *If* $q \geq 2$ *and if* $\bar{\delta} > -\frac{n}{q} + \frac{n}{2} - 1$, *then,* $W^q_{0,\bar{\delta}+m} \subset H_{0,m-1}$ *for all integer* m.

Proof This follows from the inclusion (3.7) of Appendix I, with $q = p_2$, $2 = p_1$, $\delta_2 - \delta_1 = \bar{\delta} + 1 > n(\frac{1}{2} - \frac{1}{q})$. □

Theorem 4.2 *Let* (M, γ) *be an* $M^p_{2,\delta}$ *asymptotically Euclidean manifold with* $p > \frac{n}{2}$ *and* $\delta > -\frac{n}{p}$. *Let* $a \in W^p_{0,\delta+2}$ *be given. The operator* $\triangle_\gamma - a$ *is injective on* W^q_{2,δ_q}, $\delta_q > -\frac{n}{q}$, $2 \leq q \leq p$ *if*[15]

$$\int_M \{|\partial f|^2 + af^2\}\mu_\gamma > 0 \quad \text{for any} \quad f \in W^q_{2,\delta_q}, \quad f \not\equiv 0. \tag{4.1}$$

Proof If $u \in \mathcal{D}$, space of C^∞ functions with compact support, then the following identity holds

$$\int_M u(\triangle_\gamma u - au)\mu_\gamma + \int_M (\partial u.\partial u + au^2)\mu_\gamma \equiv 0\,. \tag{4.2}$$

This identity contradicts the existence of $u \in \mathcal{D}$, $u \not\equiv 0$, with u such that $\Delta_\gamma u - au = 0$. If the identity holds also for $u \in W^q_{2,\delta_q}$, u in the kernel of $\Delta_\gamma u - au$, then it proves the injectivity of this operator on W^q_{2,δ_q}. By Lemma 3.6, such a u belongs to any $W^q_{2,\bar{\delta}}$, $\delta_q \leq \bar{\delta} < n - 2 - \frac{n}{q}$, if $n > 2$. We choose $\bar{\delta}$ such that

$$-\frac{n}{q} + \frac{n}{2} - 1 < \bar{\delta} < n - 2 - \frac{n}{q}; \tag{4.3}$$

note that the inequalities are compatible since $n > 2$. Then, by the previous lemma, $u \in W^q_{2,\bar{\delta}}$ implies that $u \in H_{0,-1}$ and $\Delta_\gamma u \in H_{0,1}$.

We approximate u by functions u_n in $\mathcal{D}$. The convergence to zero of

$$\int_M \{u\triangle_\gamma u - u_n\triangle_\gamma u_n\}\mu_\gamma \equiv \int_M \{(u - u_n)\triangle_\gamma u + u_n(\triangle_\gamma u - \Delta_\gamma u_n)\}\mu_\gamma. \tag{4.4}$$

[15] The restriction $q \geq 2$ is not necessary for the theorem to hold, but the lemma simplifies when $q \geq 2$, a condition which is satisfied in our applications.

is a simple consequence of the Cauchy–Schwarz inequality. The other limits are proved similarly. □

Corollary 4.3 *Under the same functional hypotheses on γ and a, and if $a \geq 0$, possibly $a \equiv 0$, on M, $\triangle_\gamma - a$ is injective on W^q_{2,δ_q}, $q \geq 2$, $\delta_q > -\frac{n}{q}$.*

Proof The Poincaré inequality (I.3.7) says that there exists a constant C such that

$$||f||_{H_{0,-1}} \leq C||\partial f||_{H_{0,0}}; \tag{4.5}$$

therefore the equation

$$\int_M (\partial u.\partial u + au^2)\mu_\gamma = 0 \tag{4.6}$$

which holds for all $u \in W^q_{2,\bar{\delta}}$ which lie in the kernel of $\Delta_\gamma - a$ implies, even if $a \equiv 0$, that $u \equiv 0$. □

A proof analogous to the one given in the compact case shows that if $\Delta_\gamma - a$ is injective on $W^q_{2,\delta}$ there exists a number $C_{\gamma,a}$, depending only on γ and a, such that for all $u \in W^q_{2,\delta_q}$.

$$||u||_{W^q_{2,\delta_q}} \leq C_\gamma||(\Delta_\gamma - a)u||_{W^q_{0,\delta_q+2}}. \tag{4.7}$$

The following theorem is a direct consequence of the general theorem.

Theorem 4.4 *Let (M,γ) be an $M^p_{2,\delta}$ asymptotically Euclidean manifold with $p > \frac{n}{2}$, $-\frac{n}{p} < \delta$. The Poisson operator $\Delta_\gamma - a$, $a \in L^p$, is an isomorphism from W^q_{2,δ_q} onto W^q_{0,δ_q+2} if $q \geq 2$, $-\frac{n}{q} < \delta_q < -\frac{n}{q} + n - 2$ and the inequality (4.1) is satisfied.*

Corollary 4.5 *If $\gamma \in M_{s,\delta}$, $a \in H_{s-2;\delta+2}$, $s > \frac{n}{2}$, $\frac{n}{2} - 2 > \delta > -\frac{n}{2}$, then $\Delta_\gamma - a$ is an isomorphism from $H_{s,\delta}$ onto $H_{s-2,\delta+2}$.*

4.2 *Conformal Laplace operator*

Lemma 4.6 *Let (M,γ) be an asymptotically Euclidean manifold with $\gamma \in W^p_{2,\delta}$, $p > \frac{n}{2}, \delta > -\frac{n}{p}$.*

1. *The kernel of the conformal Laplace operator $\Delta_{\gamma,conf}$, 2.45, in the space of W^q_{2,δ_q} vector fields is the space of W^q_{2,δ_q} conformal Killing vector fields, if $q \geq 2$, $-\frac{n}{q} < \delta_q < -\frac{n}{q} + n - 2$.*

2. *(M,γ) admits no conformal Killing vector field in $W^p_{2,\delta}$ under the above conditions on δ.*

Proof

1. The proof of this first result uses the identity (2.44) for smooth vector fields with compact support, and approximations analogous to those approximations made for the Poisson operator.

2. The conditions on δ imply that the considered conformal Killing fields are in a space C^0_β with $\beta > 0$, hence tend to zero at infinity. It has been proved by geometric[16] arguments, and analysis[17] arguments, that such fields do not exist on an asymptotically Euclidean manifold. □

The following theorem is a direct consequence of this lemma and Theorem 3.7.

Theorem 4.7 *Under the hypotheses of the Lemma 4.6, the conformal Laplace operator is an isomorphism from W^q_{2,δ_q} onto W^q_{0,δ_q+2}, if, in addition to its lower bound $-\frac{n}{q}$, δ_q admits the upper[18] bound $\delta_q < n - 2 - \frac{n}{q}$.*

Corollary 4.8 *If moreover $\gamma \in H_{s,\delta}$, $\delta > -\frac{n}{2}$ and $s > \frac{n}{2}$ then $\Delta_{\gamma,conf}$ is an isomorphism $H_{s,\delta} \to H_{s-2,\delta+2}$.*

5 Equation $\Delta_\gamma \varphi = f(x, \varphi)$, compact M

The Leray–Schauder theory of degree has been applied[19] to prove existence on a compact Riemannian manifold (M, γ) of solutions of an equation of the form

$$\Delta_\gamma \varphi = f(x, \varphi), \tag{5.1}$$

with application to the Lichnerowicz equation[20]. We apply here another method[21] which has the advantage to be constructive, in the sense that it is a method of successive iterations obtained by resolution of linear problems, whose bounds and convergence are obtained through the use of sub and supersolutions. We incorporate low regularity results[22] obtained recently.

To make the proofs more transparent we suppose that f is a finite sum, as it is in the Lichnerowicz equation

$$f(x, y) \equiv \sum_{I=1,\dots,N} a_I(x) y^{P_I}, \tag{5.2}$$

where the exponents P_I are given real numbers. Hence f is a smooth function of y if $y \geq \ell > 0$.

It is to solve this semi-linear equation that we have studied linear systems with $\gamma \in M^p_2$ with $p > \frac{n}{2}$, and $\gamma \in M_s$, $s > \frac{n}{2}$. The spaces W^p_2 with $p > \frac{n}{2}$, and H_s with $s > \frac{n}{2}$, are algebras, and included in C^0. If $\varphi \in W^p_2$ with $p > \frac{n}{2}$ then also $\varphi^P \in W^p_2$ [respectively $\varphi^P \in H_s$ if $\varphi \in H_s$ and $s > \frac{n}{2}$] for any number

[16] Choquet, G. and Choquet-Bruhat, Y. (1978) *C. R. Acad. Sci. Paris*, **287A**, 1047–9.

[17] Christodoulou, D. and O'Murchada, N. (1981) *Commun. Math. Phys.*, **80**, 271–300, They require somewhat more regularity.

[18] Note that this upper bound is compatible with the lower bound, because $n > 2$.

[19] Choquet-Bruhat, Y. and Leray, J. (1972) *C. R. Acad. Sci. Paris*, **274**, 81–5.

[20] Choquet-Bruhat, Y. (1972) *C. R. Acad. Sci. Paris*, **274**, 682–4.

[21] Isenberg, J. (1995) *loc. cit.*

[22] This weakening of regularity is important for the treatment of discontinuous sources, and also for the coherence with the recent developments on evolution obtained by Klainerman and Rodnianski in the case $n = 3$.

$P \geq 0$. The same result holds for a negative P if, in addition φ has no zero on the manifold M.

The bounds and convergence of successive iterations are obtained through the use of sub- and supersolutions which are defined as follows:

Definition 5.1 *A W_2^p function φ_- is called a **subsolution** of $\Delta_\gamma \varphi = f(., \varphi)$ if it satisfies on M the inequality*[23]

$$\Delta_\gamma \varphi_- \geq f(., \varphi_-). \tag{5.3}$$

*A W_2^p function φ_+ is called a **supersolution** if it satisfies*

$$\Delta_\gamma \varphi_+ \leq f(x, \varphi_+). \tag{5.4}$$

One knows by the classical maximum principle[24] that if a function $u \in C^2$ satisfies the inequality

$$\triangle_\gamma u - au \leq 0, \quad [\text{respectively} \quad \triangle_\gamma u - au \geq 0]$$

with bounded coefficients and $a \geq 0$, and if it attains a minimum $\lambda \leq 0$ [respectively a maximum $\lambda \geq 0$] at a point of M, then $u \equiv \lambda$ on[25] M. Therefore $\triangle_\gamma u - au \leq 0$ on M implies $u \geq 0$ on M if $a \geq 0$ and $a \not\equiv 0$ on M. To apply the method of sub- and supersolutions in low regularity we need the following generalized[26] maximum principle.

Lemma 5.2 *If $u \in W_2^p$ with $p > \frac{n}{2}$ satisfies the equation*

$$\Delta_\gamma u - au = -f \tag{5.5}$$

with $\gamma \in M_2^p$, $a \in L^p$, $a \geq 0$, $a \not\equiv 0$, and $f \geq 0$, then $u \geq 0$ on M.

Proof The lemma holds by the classical maximum principle if $u \in C^2$, $\gamma \in C^1$, and $a \in C^0$. We approximate $\gamma \in M_2^p$ by $\gamma_n \in M_3^p$, hence $\gamma_n \in C^1$. We approximate a by $a_n \in W_2^p \subset C^0$, $a_n \geq 0$, $a_n \not\equiv 0$ and $f \in L^p$ by $f_n \in W_2^p \subset C^0$, $f_n \geq 0$. We know by the isomorphism theorem that each equation

$$L_n \equiv \Delta_{\gamma_n} u_n - a_n u_n = -f_n$$

has one solution $u_n \in W_4^p \subset C^2$, and $u_n \geq 0$.

The difference $v_n \equiv u - u_n$ satisfies the equation:

$$\Delta_\gamma v_n - a v_n = F_n, \tag{5.6}$$

[23] This inequality is intended to hold almost everywhere in the L^p sense.

[24] Protter, M. H. and Weinberger, H. F. (1967) *Maximum Principles in Differential Equations*, Prentice Hall, p. 64.

[25] Also if S is a bounded domain of M with smooth boundary a minimum ≤ 0 [respectively a maximum ≥ 0] must be attained on the boundary.

[26] Various such generalizations appear in the literature, but to remain relatively self-contained we prefer to give our own formulation and proof.

with

$$F_n \equiv (f_n - f) + (\Delta_\gamma - \Delta_{\gamma_n})u_n + (a - a_n)u_n. \tag{5.7}$$

We know (Theorem 2.3) that since $L \equiv \Delta_\gamma - a$ is injective on W_2^p there exists a number C_L depending only on γ and a such that

$$||v_n||_{W_2^p} \leq C_L ||F_n||_{L^p}. \tag{5.8}$$

We know that, if $p > \frac{n}{2}$, then $W_2^p \times L^p$ and $W_1^p \times W_1^p$ are continuously embedded in L^p. The expression of F_n implies that

$$||F_n||_{L^p} \leq ||f - f_n||_{L^p} + \{||\gamma - \gamma_n||_{W_2^p} + ||a - a_n||_{L^p}\}||u||_{W_2^p}. \tag{5.9}$$

We deduce from this inequality that the sequence v_n converges to zero in W_2^p, and consequently u_n converges to $u \in W_2^p$, so long as f_n converges to f and γ_n to γ. Since $u_n \geq 0$, we have $u \geq 0$, everywhere, because $W_2^p \subset C^0$ when $p > \frac{n}{2}$. □

We can now prove the following theorem.

Theorem 5.3 *(Existence) Equation (5.2), $\Delta_\gamma \varphi = f(., \varphi)$, on the compact manifold (M, γ) admits a solution $\varphi \in W_2^p$, $p > \frac{n}{2}$, so long as the following conditions hold:*

(a) *$\gamma \in M_2^p$, $p > \frac{n}{2}$, and $a_I \in L^p$, I=1,...,N.*

(b) *The equation admits a subsolution φ_- and a supersolution φ_+, both in W_2^p, $0 < \ell \leq \varphi_- \leq \varphi_+ \leq m$.*

The solution is such that $\varphi_- \leq \varphi \leq \varphi_+$.
This solution is unique if $f(., y)$ is monotonically increasing in y.

Proof The proof follows the same lines as the proof given in Isenberg (1995), but uses the generalized maximum principle. We define successive iterates by a solution of the following elliptic equation, linear in φ_n so long as φ_{n-1} is known,

$$\Delta_\gamma \varphi_n - a\varphi_n = f(., \varphi_{n-1}) - a\varphi_{n-1}, \tag{5.10}$$

with $a \in L^p$, $a \geq k$, k a strictly positive number to be determined later, introduced now to allow the application of the isomorphism theorem 2.7.

We take $\varphi_0 = \varphi_-$ to start the iteration; i.e. we define φ_1 by solving the equation

$$\Delta_\gamma \varphi_1 - a\varphi_1 = f(., \varphi_-) - a\varphi_-. \tag{5.11}$$

It follows from the hypothesis made on γ, a, f and φ_- that such a solution φ_1 exists and is unique in W_2^p. Moreover we have by the definitions of φ_1 and φ_- that

$$\Delta_\gamma(\varphi_1 - \varphi_-) - a(\varphi_1 - \varphi_-) \leq 0 \tag{5.12}$$

The conclusion $\varphi_1 - \varphi_- \geq 0$ on M follows from the generalized maximum principle.

We now use the definition of φ_+ to obtain

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) \leq f(x,\varphi_+) - f(x,\varphi_-) - a(\varphi_+ - \varphi_-) \tag{5.13}$$

from which we derive, using the mean value theorem (CB-DM1, p. 78)

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) \leq (\varphi_+ - \varphi_-) \int_0^1 \{f'_\varphi(x,\varphi_- + t(\varphi_+ - \varphi_-)) - a\} dt \tag{5.14}$$

If we now choose the function a so that

$$a = Sup_{l \leq y \leq m} |f'_y(.,y)| + k, \quad k > 0, \tag{5.15}$$

then we have on M that $a \geq k > 0$ and

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) = F \leq 0, \tag{5.16}$$

since on M it holds that $0 < \ell \leq \varphi_- + t(\varphi_+ - \varphi_-) \leq m$. For example, if the function f is given by (5.2), then

$$\begin{aligned} |f'_y(.,y)| &\leq \sum_{I=1,\ldots N} |a_I P_I| y^{P_I - 1} \\ &\leq \sum_{I_+} |a_{I_+} P_{I_+}| m^{P_{I_+} - 1} + \sum_{I_-} |a_{I_-} P_{I_-}| \ell^{P_{I_-} - 1} + |a_{I_0}| \end{aligned} \tag{5.17}$$

with $P_{I_+} > 1$, $P_{I_-} < 1$, $P_{I_0} = 1$. Under the hypothesis made on a_I the maximum principle holds; hence $\varphi_1 \leq \varphi_+$ on M.

Suppose that there exists $\varphi_m \in W_2^p$, $m = 1, \ldots, n-1$, such that on M

$$\varphi_- \leq \varphi_m \leq \varphi_+. \tag{5.18}$$

Then φ_n exists on M as a solution in W_2^p of 5.10. Suppose moreover that

$$\varphi_{n-1} - \varphi_{n-2} \geq 0. \tag{5.19}$$

Then the equations satisfied by φ_n and φ_{n-1} imply that

$$\Delta_\gamma(\varphi_n - \varphi_{n-1}) - a(\varphi_n - \varphi_{n-1}) = (b_n - a)(\varphi_{n-1} - \varphi_{n-2}) \tag{5.20}$$

where b_n is the function in L^p

$$b_n \equiv \int_0^1 f'_\varphi(., \varphi_{n-2} + t(\varphi_{n-1} - \varphi_{n-2}) dt. \tag{5.21}$$

It results from the definition of a that

$$b_n - a \leq -k < 0. \tag{5.22}$$

This inequality together with $\varphi_{n-1} \geq \varphi_{n-2}$ ensures that the right-hand side of (5.20) is negative; hence $\varphi_- \leq \varphi_{n-1} \leq \varphi_n$. To show that the sequence is bounded above by φ_+ we write again an inequality of the form

$$\begin{aligned} \Delta_\gamma(\varphi_+ - \varphi_n) - a(\varphi_+ - \varphi_n) &\leq f(\varphi_+) - f(\varphi_{n-1}) - a(\varphi_+ - \varphi_{n-1}) \\ &\leq (b_n - a)(\varphi_+ - \varphi_{n-1}) \leq 0, \end{aligned} \tag{5.23}$$

from which the conclusion follows. We have proved the existence of the sequence $\varphi_n \in W_2^p$, with $0 \leq \varphi_- \leq \varphi_n \leq \varphi_+$; hence $||\varphi_n||_{L^p}$ is uniformly bounded by $||\varphi_+||_{L^p}$. Due to the defining equation (5.10), $||\varphi_n||_{W_2^p}$ is also uniformly bounded. Since functions in W_2^p are equicontinuous, we can extract (Ascoli–Arzela theorem) from the sequence φ_n a subsequence, $\tilde{\varphi}_n$, which converges in C^0 norm, to a function $\varphi \in C^0$. On the other hand the sequence φ_n of continuous and positive functions is increasing and bounded above by φ_+; hence it is pointwise convergent to a function φ, which coincides with the previously obtained φ. The whole sequence (not only a subsequence) is therefore convergent.

We show that the limit is in fact in W_2^p by using the following elliptic estimate applied to Equation (5.20):

$$\| \varphi_n - \varphi_{n-1} \|_{W_2^p} \leq C_{\gamma,a} ||(b_n - a)(\varphi_{n-1} - \varphi_{n-2}) \|_{L^p} \tag{5.24}$$

which implies that

$$\| \varphi_n - \varphi_{n-1} \|_{W_2^p} \leq C_{\gamma,a} ||(b_n - a)||_{L^p} ||\varphi_{n-1} - \varphi_{n-2}) \|_{C^0} \, . \tag{5.25}$$

The sequence φ_n converges in W_2^p (to φ, by uniqueness of limits) since it converges in C^0, and the $b_n - a$ are uniformly bounded in L^p norm. The limit satisfies Equation (5.1) in the L^p sense.

We remark that the solution we have constructed depends on the choice of the initial φ_0. We have taken φ_- as an initial φ_0. We could have taken φ_+ and constructed a decreasing and bounded below sequence which converges to a limit Ψ. We cannot in the procedure used here start from an arbitrary φ_0, with $\varphi_- \leq \varphi_0 \leq \varphi_+$, because then we would not know if φ_1 satisfies the same inequalities.

The uniqueness in the case $f'_y > 0$ is a consequence of the equation satisfied by the difference of two solutions, namely

$$\Delta_\gamma(\varphi - \psi) - \left\{ \int_0^1 f'_\varphi(\psi + t(\varphi - \psi) dt \right\} (\varphi - \psi) = 0. \qquad \square$$

A more general uniqueness theorem holds for the Lichnerowicz equation; see Chapter 7.

Corollary 5.4 *The theorem still holds if in hypotheses and conclusions the spaces M_2^p, W_2^p and L^p are replaced respectively by M_s, H_s and H_{s-2}, $s > \frac{n}{2}$.*

The proof of the corollary rests on Corollary 2.8.

6 $\Delta_\gamma\varphi = f(x,\varphi)$ *on* (M,γ) asymptotically Euclidean

We take again f of the form (5.2), where the a_Is are functions on M in $W^p_{0,\delta+2}$.Then the function on M given by $x \mapsto f(x,\varphi(x))$, denoted for convenience $f(x,\varphi)$, belongs to $W^p_{0,\delta+2}$ if φ is continuous and it is bounded on M and bounded away from zero in the case that some of the exponents P_I are negative.

A subsolution φ_- and a supersolution φ_+ of $\triangle_\gamma\varphi = f(x,\varphi)$ satisfy the inequalities (5.3) and (5.4) given in the case of a compact manifold. We will specify the hypotheses made on these functions in the theorem that we will prove, but first we prove an adapted version of the generalized maximum principle, stated in the form that we use here.

We say that a continuous function f tends to a value $c \in R$ at infinity if for any $\varepsilon \geq 0$ there exists a compact S such that:

$$\sup_{M-S} |f - c| \leq \varepsilon. \tag{6.1}$$

Lemma 6.1 *(Generalized maximum principle) Let ψ satisfy the equation*

$$L\psi := \Delta_\gamma\psi - a\psi = -f \tag{6.2}$$

with $\gamma \in M^p_{2,\delta}$, $p > \frac{n}{2}$, $-\frac{n}{p} < \delta$, $a \in W^p_{0,\delta+2}$, $a \geq 0$, and $f \geq 0$. Suppose that $\psi - c \in W^p_{2,\delta}$, with $c \geq 0$ a given number. Then $\psi \geq 0$ on M.

Proof One knows by the classical maximum principle that if a solution $\psi \in C^2$ of the inequality with bounded coefficients

$$\triangle_\gamma\psi - a\psi \leq 0, \quad a \geq 0 \tag{6.3}$$

attains a minimum $\lambda \leq 0$ at a point of M then $f \equiv \lambda$ on M. Also, if S is a bounded domain of M with smooth boundary ∂S where ψ satisfies this inequality and if ψ attains a minimum ≤ 0 in S, then this minimum must be attained on the boundary ∂S [27]. Hence if ψ tends to $c \geq 0$ at infinity it satisfies $\psi \geq 0$ on M. Indeed, suppose that ψ takes a negative value α at some point $x \in M$. Choose $\varepsilon < |\alpha|$. By the hypothesis on ψ there is a compact S such that

$$\sup_{M-S} |\psi - c| \leq \varepsilon\ . \tag{6.4}$$

Take a relatively compact open set D containing S and x. If ψ attains a minimum ≤ 0, it is on the boundary ∂D; i.e. in $M - S$. This contradicts the fact that the absolute value of this minimum is necessarily greater than or equal to $|\alpha|$, itself greater than ε, which is the maximum of $|\psi|$ in $M - S$, a contradiction. An analogous proof shows that a C^2 solution ψ of (6.3) tending to $c \leq 0$ at infinity is such that $\psi \leq 0$ on M.

The theorem therefore holds if $\gamma \in M^p_{3,\delta}$, $a \in W^p_{2,\delta+2}$, $a \geq 0$, $\psi - c \in W^p_{4,\delta}$ $\delta > -\frac{n}{p}$, since the coefficients of (5.3) are then bounded and $\psi - c \in C^2_\beta$, $\beta > 0$. We

[27] See Protter and Weinberger (1967), p. 64.

approximate $\gamma \in M^p_{2,\delta}$ by $\gamma_n \in M^p_{3,\delta}$, a by $a_n \in W^p_{2,\delta+2}$, $a_n \geq 0$ and $f \in W^p_{0,\delta+2}$ by $f_n \in W^p_{2,\delta+2}$, $f_n \geq 0$. We define $\psi_n \equiv c + u_n$ by solving the equations

$$\Delta_{\gamma_n} u_n - a_n u_n = -f_n + ca_n$$

These solutions $u_n \in W^p_{4,\delta} \subset C^2_\beta$ exist by the isomorphism theorem. The corresponding functions ψ_n tend to c at infinity and satisfy the equations

$$\Delta_{\gamma_n} \psi_n - a_n \psi_n = -f_n. \tag{6.5}$$

They are non-negative by the classical maximum principle. We show that ψ_n converges to the solution ψ of the given equation, as we did in the compact case. The difference $v_n \equiv \psi - \psi_n$ belongs to $W^p_{2,\delta}$ and satisfies the equation

$$\triangle_\gamma \psi - a\psi = -f, \quad a \geq 0 \tag{6.6}$$

$$\Delta_\gamma v_n - a v_n = F_n. \tag{6.7}$$

with

$$F \equiv (f_n - f) + (\Delta_\gamma - \Delta_{\gamma_n})\psi_n + (a - a_n)\psi_n. \tag{6.8}$$

The isomorphism theorem implies that

$$||v_n||_{W^p_{2,\delta}} \leq C_{\gamma,a} ||F_n||_{W^p_{0,\delta+2}}. \tag{6.9}$$

The convergence to zero of the right-hand side results from the hypotheses, therefore v_n converges to zero in $W^p_{2,\delta+2}$. The convergence is pointwise if $p > \frac{n}{2}$. □

We can now prove our fundamental theorem.

Theorem 6.2 *Suppose the equation $\Delta_\gamma \varphi = f(x, \varphi)$, with $\gamma \in M^p_{2,\delta}$, $p > \frac{n}{2}$, $\delta > -\frac{n}{p}$, f of the type (5.2) admits sub- and supersolutions $\varphi_- \leq \varphi_+$ such that:*

$$\varphi_- - c_-,\ \varphi_+ - c_+ \in W^p_{2,\delta} \tag{6.10}$$

with c_- and c_+ numbers which satisfy

$$0 \leq c_- \leq 1 \leq c_+. \tag{6.11}$$

Suppose also that $Inf \varphi_- > 0$ if some of the exponents P_I are negative. Then the equation admits a solution φ such that

$$\varphi_- \leq \varphi \leq \varphi_+\ ,\ \ 1 - \varphi \in W^p_{2,\delta} \quad \text{if } -\frac{n}{p} < \delta < n - 2 - \frac{n}{p}\ . \tag{6.12}$$

If moreover $\gamma \in M^p_{2+s}$ and $a_I \in W^p_{s,\delta+2}$, then the solution $\varphi = 1 + u$ has $u \in W^p_{s+2,\delta}$

Proof It is analogous to the one given in the case of a compact manifold: we construct a solution by induction, starting from φ_-. We define $u_1 \equiv \varphi_1 - 1 \in W^p_{2,\delta}$

as the solution of the equation

$$\Delta_\gamma u_1 - a u_1 = f(x,\varphi_-) - a(\varphi_- - 1) \tag{6.13}$$

with $a \in W^p_{0,\delta+2}$ a positive function on M, to be chosen later

The difference $\psi_1 \equiv \varphi_1 - \varphi_-$ tends to $1 - c \geq 0$ at infinity and satisfies the inequality

$$\Delta_\gamma \psi_1 - a\psi_1 = -\Delta_\gamma \varphi_- + f(x,\varphi_-) \leq 0. \tag{6.14}$$

The generalized maximum principle tells us that

$$\psi_1 \geq 0; \quad \text{i.e.} \quad \varphi_1 \geq \varphi_-. \tag{6.15}$$

To obtain that $\varphi_1 \leq \varphi_+$ we choose a such that almost everywhere on M we have

$$a \geq \sup_{\ell \leq y \leq m} |f'_y(.,y)|, \quad |f'_y(.,y)| \leq |a_P| \left| \frac{\partial b_P(y)}{\partial y} \right|. \tag{6.16}$$

Such a function $a \in W^p_{0,\delta+2}$ exists by the hypothesis made on f.

We now use the definition of φ_+ to obtain

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) \leq f(x,\varphi_+) - f(x,\varphi_-) - a(\varphi_+ - \varphi_-), \tag{6.17}$$

from which we deduce, using the mean value theorem, that

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) \leq (\varphi_+ - \varphi_-) \int_0^1 \{f'_y(x,\varphi_- + t(\varphi_+ - \varphi_-)) - a\} dt \tag{6.18}$$

We see that, as a consequence of the choice of the function a, we have on M, as in the compact case

$$\Delta_\gamma(\varphi_+ - \varphi_1) - a(\varphi_+ - \varphi_1) \leq 0. \tag{6.19}$$

Hence $\varphi_+ \geq \varphi_1$ if $c_+ \geq 1$.

We determine $u_n \equiv \varphi_n - 1$ by the induction formula

$$\triangle_\gamma u_n - a u_n = f(x,\varphi_{n-1}) - a u_{n-1} \tag{6.20}$$

and we obtain, as in the compact case, a nondecreasing sequence of continuous functions bounded above and below

$$\varphi_- \leq \varphi_{n-1} \leq \varphi_n \leq \varphi_+. \tag{6.21}$$

The sequence of continuous functions φ_n therefore converges at each point $x \in M$ to a limit $\varphi(x) = 1 + u(x)$, with $\varphi_- \leq \varphi \leq \varphi_+$.

To show that $\varphi - 1 \in W^p_{2,\delta}$ and that φ is a solution of the given equation, we use the fact that the functions $f(x,\varphi_n) - a u_n$ belong to $W^p_{0,\delta+2}$ with uniformly bounded norms. The linear elliptic inequality

$$\|u_{n+1}\|_{W^p_{2,\delta}} \leq C \|f(x,\varphi_n) - a u_n)\|_{W^p_{0,\delta+2}} \tag{6.22}$$

shows that the sequence u_n is uniformly bounded in the $W^p_{2,\delta} \subset C^{0,\alpha}_\beta$ norm, $\beta > 0$, because $p > \frac{n}{2}$ and $\delta > -\frac{n}{p}$. Since $C^{0,\alpha}_\beta$ is compactly embedded in $C^0_0 \equiv \bar{C}^0$ (continuous and bounded functions) there is a subsequence, still denoted u_n, which converges in $\bar{C}^0$ norm to a function $u \in \bar{C}^0$, identical to the previously found u (to which the whole sequence therefore converges).

The functions $f(x, \varphi_n)$ converge to $f(x, \varphi)$ in the $W^p_{0,\delta+2}$ norm as a consequence of the trivial inequality

$$\|f(x,\varphi) - f(x,\varphi_n)\|_{W^p_{0,\delta+2}} \leq C\|\varphi - \varphi_n\|_{C^0_0}\|a\|_{W^p_{0,\delta+2}}\,, \tag{6.23}$$

where C depends only on (M, e) and a is defined by (6.16). It results therefore from the elliptic inequality that u_n converges to u in $W^p_{2,\delta}$ norm, and that $\varphi = 1 + u$, $\varphi_- \leq \varphi \leq \varphi_+$ satisfies the given equation.

The regularity $u \in W^p_{s+2,\delta}$ for more regular coefficients is proved by induction, after checking that the hypotheses on a_I imply that the function $x \mapsto f(x, \varphi(x))$ belongs to $W^p_{s',\delta+2}$, $0 \leq s' \leq s$ so long as $\varphi - 1 \in W^p_{s'+2,\delta}$. This result is proved by using the calculus derivation formulas and the multiplication properties of weighted Sobolev spaces. □

Corollary 6.3 *The theorem still holds if in the hypotheses and the conclusions the spaces $M^p_{2,\delta}$, $W^p_{2,\delta}$ and $W^p_{0,\delta}$ are replaced respectively by $M_{s,\delta}$, $H_{s,\delta}$, and $H_{s-2,\delta}$, $s > \frac{n}{2}$ and $-\frac{n}{2} < \delta < -2 + \frac{n}{2}$.*

Proof This rests on Corollary 4.5. □

APPENDIX III

QUASI-DIAGONAL, QUASI-LINEAR, SECOND-ORDER HYPERBOLIC SYSTEMS

1 Introduction

In this appendix we prove local in time existence and uniqueness theorems for solutions of the Cauchy problem on a manifold for quasilinear systems of second-order partial differential equations whose principal second-order terms are diagonal: the Einstein equations in harmonic gauge are of this type, as well as many other field equations. These systems are a particular case of Leray hyperbolic systems, but the energy estimates that we will use[1] to prove the existence of solutions are much easier to obtain than in the general case. We use simply the Sobolev spaces H_s on space manifolds with integer s and we require the principal coefficients, i.e. the Lorentzian metric, to be C^1. More refined results in progress involve more refined techniques[2], outside of the scope of this book.

We begin with the wave equation on a Lorentzian sliced manifold ($V = M \times R, \mathbf{g}$). For such an equation the energy estimate has a clear geometric and physical meaning. We consider in subsequent sections linear second-order quasi-diagonal systems with a wave-like equation as principal part. The last sections treat semilinear and quasilinear systems. The proof of the local in time existence theorem of solutions of these semilinear or quasilinear systems uses iteration of solutions of linear systems, some fixed point theorems, and, in the case of quasilinear systems, some functional analysis arguments.

We give general methods and results, with proofs which may help readers to find for themselves the variants that they may need in particular circumstances.

2 Wave equation on (V,$\mathbf{g}$)

We consider the wave equation with source f and unknown a scalar function u,

$$\Box_{\mathbf{g}} u \equiv g^{\alpha\beta}\nabla_\alpha \partial_\beta u = f, \tag{2.1}$$

[1] The solutions can also be obtained constructively through parametrization (Fourès (Choquet)-Bruhat, Y. (1952) *Acta Matematica*, **88**, 141–225.

[2] Recent developments using elaborate methods of Fourier analysis relax somewhat the hypothesis on the regularity required from the data (Klainerman, S. and Rodnianski, I. (2007) *J. Hyp. Diff. Eq.*, **4**(3), 401–33, arXiv:math.AP/0109173 vl).

on a sliced Lorentzian manifold[3] $(V, \mathbf{g})$; ∇_α is the covariant derivative in the metric $\mathbf{g}$; the source f is a given function on V.

2.1 *Definitions*

We recall definitions and notations used in previous chapters.

Definition 2.1 *A pair $(V, \mathbf{g})$ is a sliced Lorentzian (s.L.) manifold if*

1. *$V = M \times R$, M a smooth n-dimensional manifold which we suppose oriented. V is then also oriented and time oriented by the increasing direction on R.*

2. *The spacetime metric $\mathbf{g}$ is pseudo-Riemannian with signature $(-, +, \ldots, +)$, and each submanifold $M_t \equiv M \times \{t\}$ is spacelike. The lines $\{x\} \times R$ are transversal to them.*

In a Cauchy adapted frame (see Section VI.3) we set

$$\theta^0 := dt, \quad \theta^i = dx^i + \beta^i dt \tag{2.2}$$

where the x^i are local coordinates on M. The metric on an s.L. manifold can be written

$$\mathbf{g} = -N^2 dt^2 + g_{ij}\theta^i\theta^j. \tag{2.3}$$

The metrics

$$g_t = g_{ij}(., t)dx^i dx^j \tag{2.4}$$

induced by $\mathbf{g}$ on M_t are (properly) Riemannian. The lapse N and the shift β are the projection of the tangent vector to the transversal lines respectively on the normal and on the tangent plane to M_t. The lapse N is strictly positive.

We denote by n the past-oriented unit normal to M_t, i.e. such that

$$n_0 = N, \quad n^0 = -N^{-1}. \tag{2.5}$$

We denote by ∂_0, ∂_i the Pfaff derivatives in the coframe dt, θ^i, namely:

$$\partial_i \equiv \partial/dx^i, \quad \partial_0 \equiv \partial/\partial t - \beta^i \partial/\partial x^i. \tag{2.6}$$

2.2 *Stress energy tensor. Energy momentum vector*

The stress energy tensor associated with u is (indices raised with $\mathbf{g}$)

$$T_{\alpha\beta} \equiv \partial_\alpha u \partial_\beta u - \frac{1}{2} g_{\alpha\beta} \partial_\lambda u \partial^\lambda u \tag{2.7}$$

Lemma 2.2 *When u is a solution of the wave equation with source f the $\mathbf{g}$-divergence of its stress energy tensor is such that*

$$\nabla_\alpha T^{\alpha\beta} \equiv f \partial^\beta u \tag{2.8}$$

[3] We don't use bold face for the spacetime metric when Greek indices are used.

Proof Computation shows that, changing names of indices

$$\nabla_\alpha T^{\alpha\beta} \equiv (\nabla_\alpha \partial^\alpha u)\partial^\beta u + \partial^\alpha u(\nabla_\alpha \partial^\beta u - \nabla^\beta \partial_\alpha u),$$

the announced result follows when u is a scalar function which satisfies the wave equation. □

Definition 2.3 *The energy momentum vector $\mathcal{P}$ associated with u and a vector field X is defined by:*

$$\mathcal{P}^\alpha := T^{\alpha\beta} X_\beta. \tag{2.9}$$

Lemma 2.4 *If X is timelike or null, then $\mathcal{P}$ is timelike or null.*

Proof We have, with a dot denoting the scalar product in **g**

$$\mathcal{P} \equiv \partial u(X.\partial u) - \frac{1}{2} X(\partial u.\partial u); \tag{2.10}$$

therefore by elementary computation

$$\mathcal{P}.\mathcal{P} \equiv \frac{1}{4}(X.X)(\partial u.\partial u)^2 \leq 0 \quad \text{if} \quad X.X \leq 0. \qquad \square$$

The vectors $\mathcal{P}$ and X have opposite time orientation, as shown by the following definition and lemma.

Definition 2.5 *We denote by $G_{X,\ell}$ the quadratic form in ∂u defined by:*

$$\frac{1}{2} G_{X,\ell}(\partial u, \partial u) := \mathcal{P}.\ell. \tag{2.11}$$

The definition of $\mathcal{P}$ and the expression of T gives

$$\frac{1}{2} G^{\alpha\beta}_{X,\ell} \partial_\alpha u \partial_\beta u := \mathcal{P}.\ell \equiv \mathcal{P}^\alpha \ell_\alpha \equiv \frac{1}{2} T^{\alpha\beta}(X_\beta \ell_\alpha + X_\alpha \ell_\beta);$$

that is,

$$G^{\alpha\beta}_{X,\ell} \equiv X^\alpha \ell^\beta + X^\beta \ell^\alpha - X^\lambda \ell_\lambda g^{\alpha\beta}. \tag{2.12}$$

If we choose $X = n$ then $G_{n,n}$ is the positive definite quadratic form on covariant vectors ξ:

$$G_{n,n}(\xi, \xi) \equiv N^{-2}\xi_0^2 + g^{ij}\xi_i\xi_j. \tag{2.13}$$

More generally

Lemma 2.6 *At each point of V where X is timelike and ℓ timelike or null, with the same time orientation than X, the quadratic form $G_{X,\ell}$ is positive. It is positive definite if ℓ is restricted to be timelike.*

Proof To show that $G_{X,\ell}$ is positive at any point of V we choose at that point an orthonormal frame in the tangent space to V with time axis in the direction

of X and space axis orthogonal to it. In this frame the components of X are $X^0 = -X_0$, $X^i = X_i = 0$, while $\ell_0 = -\ell^0$. The quadratic form G reads then, for any covariant vector ξ,

$$G^{\alpha\beta}_{X,\ell}\xi_\alpha\xi_\beta = X^0\ell^0\left\{(\xi_0)^2 + 2\ell^i(\ell^0)^{-1}\xi_i\xi_0 + \sum_i(\xi_i)^2\right\}. \tag{2.14}$$

If ℓ is timelike or null it satisfies in the orthonormal frame the inequality:

$$\sum_i \mid \ell^i \mid^2 \leq \mid \ell^0 \mid^2$$

The Schwartz inequality gives then:

$$|\ell^i\xi_i| \leq |\ell^0|\left\{\sum_i \mid \xi_i|^2\right\}^{\frac{1}{2}}$$

The quantity between parentheses in (2.14) is therefore greater or equal to the square of $\{\sum_i \mid \xi_i \mid^2\}^{\frac{1}{2}} - \ell^0$ with strict inequality if the inequality for ℓ is strict. The quadratic form $G_{X,\ell}(\xi,\xi)$ is therefore positive or zero if X and ℓ have the same time orientation; it is strictly positive if ℓ is timelike. □

It results from the previous lemma that the matrix associated with the quadratic form $G_{X,\ell}$ is invertible if ℓ is a timelike vector ν. The inverse $(G_{X,\nu})^{-1}$ is then also positive definite.

Exercise. Set $\ell = n$, the unit normal vector to M_t. Show that $G^{-1}_{X,n}$ is the symmetric covariant 2-tensor

$$(G^{-1}_{X,n})_{\alpha\beta} \equiv \frac{1}{\gamma}\left(g_{\alpha\beta} + n_\alpha n_\beta + \frac{1}{\alpha^2}X_\alpha X_\beta\right) \tag{2.15}$$

with

$$\mathbf{g}(X,X) \equiv X^\alpha X_\alpha = -\alpha^2, \quad \mathbf{g}(X,n) = -\gamma. \tag{2.16}$$

Remark. one always have $\alpha^2 \leq \gamma^2$ because of the scalar product inequality for timelike vectors on a Lorentzian manifold

$$|\mathbf{g}(X,n)|^2 \geq |\mathbf{g}(X,X)||g(n,n)|.$$

2.3 *Energy density*

We denote by $\varepsilon_{X,t}(u)$ the square of the $G_{X,n}$ norm of the gradient ∂u of u; it is the density of energy with respect to X of the field u at a point of the s.L. manifold V at time t,

$$\varepsilon_{X,t}(u) := \mathcal{P}^\alpha n_\alpha \equiv \frac{1}{2} \mid \partial u \mid^2_{G_{X,n}} \equiv \frac{1}{2}G_{X,n}(\partial u, \partial u) \equiv \frac{1}{2}G^{\alpha\beta}_{X,n}\partial_\alpha u\partial_\beta u.$$

This density of energy depends on the choice of the vector X. It is positive if X is timelike or null and past oriented.

Examples.

- Take for X the vector n itself, that is in the chosen coframe (dt, θ^i):

$$X_0 = N, \quad X_i = 0, \quad \text{hence} \quad X^0 = -N^{-1}, \quad X^i = 0.$$

This choice of X is called the **canonical choice**. It is natural when the metric **g** does not possess a privileged, for instance Killing, timelike vector field. We denote simply by $\varepsilon_t(u)$ the corresponding energy density. It is given by the positive definite quadratic form:

$$\varepsilon_t(u) \equiv \frac{1}{2} G_{n,n}(\partial u, \partial u) \equiv \frac{1}{2}\{(N^{-1}\partial_0 u)^2 + g^{ij}\partial_i u \partial_j u\} \tag{2.17}$$

(recall that ∂_α denote Pfaff derivatives in the Cauchy adapted coframe, given by (2.2), g^{ij} is the positive definite contravariant associate to g_{ij}).

- Take for X the past-oriented tangent vector to the transversal lines $\{x\} \times R$. It has for components in our Cauchy-adapted coframe:

$$X^0 = -1, \quad X^i = -\beta^i$$

Then we have:

$$\varepsilon_{X,t}(u) = \frac{1}{2}\{N^{-2}(\partial_0 u)^2 - 2N^{-1}\beta^i \partial_0 u \partial_i u + g^{ij}\partial_i u \partial_j u\} \tag{2.18}$$

We check, as announced, that this $\varepsilon_{X,t}(u)$ is a positive definite quadratic form in ∂u if our chosen X is timelike, since we have then:

$$N^{-2} g_{ij}\beta^i \beta^j < 1.$$

Such a choice of X is particularly appropriate when **g** has a timelike Killing vector field whose trajectories are chosen to obtain the factorization $V = M \times R$.

2.4 *Energy equality on a compact domain*

The ***g***-divergence of the energy momentum vector $\mathcal{P}$ is (we use the symmetry of the tensor $T^{\alpha\beta}$)

$$\nabla_\alpha \mathcal{P}^\alpha \equiv X_\beta \nabla_\alpha T^{\alpha\beta} + \frac{1}{2} T^{\alpha\beta}(\nabla_\alpha X_\beta + \nabla_\beta X_\alpha). \tag{2.19}$$

Hence, if u is a solution of the wave equation with source f, it holds that:

$$\nabla_\alpha \mathcal{P}^\alpha = fX.\partial u + \frac{1}{2} T.\mathcal{L}_X(\mathbf{g}), \tag{2.20}$$

where $\mathcal{L}_X(\mathbf{g})$ is the Lie derivative of **g** with respect to the vector field X.

Remark 2.7 *With the canonical choice $X=n$ it holds that (see Section VI.3.4)*

$$(\mathcal{L}_n\mathbf{g})_{00} = 0, \quad (\mathcal{L}_n\mathbf{g})_{i0} = \partial_i N, \quad (\mathcal{L}_n\mathbf{g})_{ij} = 2K_{ij}. \tag{2.21}$$

In order to integrate the divergence formula (2.20) using elementary calculus we make the following regularity hypothesis on $\mathbf{g}$, f, and X:

We suppose that the spacetime metric $\mathbf{g}$ is C^1. The energy momentum vector $\mathcal{P}$ is then C^1 if the vector field X is C^1 and u is C^2. We suppose that the source f is C^0. The right-hand side of the divergence equation (2.20) for $\mathcal{P}$ is then also C^0.

We denote by $\omega_{\mathbf{g}}$ the $n+1$ volume form of $\mathbf{g}$; it reads in local coordinates[4]

$$\omega_{\mathbf{g}} = |\det \mathbf{g}|^{\frac{1}{2}} dx^0 \wedge dx^1 \cdots \wedge dx^n. \tag{2.22}$$

We recall the identity (d denotes the exterior derivative)

$$\nabla_\alpha \mathcal{P}^\alpha \omega_{\mathbf{g}} := \partial_\alpha(|\det \mathbf{g}|^{\frac{1}{2}} \mathcal{P}^\alpha) dx^0 \wedge dx^1 \cdots \wedge dx^n \equiv d(\mathcal{P}.\omega_L) \equiv d(\mathcal{P}^\alpha \bar{\omega}_\alpha) \tag{2.23}$$

where $\bar{\omega}$ is the covariant vector valued Leray n form whose components are given by

$$\bar{\omega}_\alpha = (-1)^\alpha |\det g|^{\frac{1}{2}} dx^0 \wedge dx^1 \cdots \wedge d\hat{x}^\alpha \wedge \cdots \wedge dx^n$$

the notation $\hat{\alpha}$ means that the corresponding differential does not appear in the n-form $\bar{\omega}_\alpha$.

Definition 2.8 *A domain (simply connected open subset) Ω of V, with boundary $\partial\Omega$, is called* ***Stokes regular*** *if the Stokes formula,*

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega, \text{ with } \omega \text{ any smooth } n\text{-form}, \tag{2.24}$$

holds; for instance (cf. Choquet-Bruhat (1968)[5]) the boundary $\partial\Omega$ is an n-dimensional manifold which is C^1 by pieces[6].

The integration with respect to the volume form $\omega_{\mathbf{g}}$ of Equation (2.23) giving $\nabla_\alpha \mathcal{P}^\alpha$ in a Stokes regular domain Ω gives the geometric integral equality

$$\int_\Omega \left\{ fX.\partial u + \frac{1}{2} T.\mathcal{L}_X g \right\} \omega_{\mathbf{g}} = \int_{\partial\Omega} \mathcal{P}^\alpha.\bar{\omega}_\alpha, \tag{2.25}$$

where the boundary $\partial\Omega$ has the orientation induced by the orientation of Ω, deduced itself from the orientation of V, i.e. the orientation of M and the natural orientation of R (increasing t).

Consider a differentiable subset U of $\partial\Omega$ with equation $t + F(x) = 0$, $x \in \bar{U} \subset M$. On this subset it holds that $dx^0 + \partial_i F dx^i = 0$. Using

$$\det \mathbf{g} = N \det g \tag{2.26}$$

[4] We denote indifferently by t or x^0 the time variable.

[5] Choquet-Bruhat, Y. (1968) *Géométrie Différentielle et Systèmes Extérieurs*, Dunod, p. 78.

[6] For the less stringent hypothesis implying the validity of Stokes formula cf. Whitney, H. *Integration theory*, Princeton University Press.

we find that, on $U \subset \partial\Omega$

$$\bar{\omega}_\alpha = \ell_\alpha \omega_g \quad \text{with} \quad \ell_0 = N, \quad \ell_i = N\partial_i F$$

where ω_g is the volume form of the space metric g,

$$\omega_g := |\det g|^{\frac{1}{2}} dx^1 \wedge \cdots \wedge dx^n /$$

Remark 2.9 *The vector ℓ appearing in the above formula is proportional to the gradient of the equation of $\partial\Omega$ and past oriented, since it satisfies $\ell_0 = N$, i.e. $\ell^0 = -N^{-1} < 0$.*

Let U be included in a slice M_t; then

$$\bar{\omega}_0 = N\bar{\omega}_g \equiv \omega_{\bar{g}}, \quad \bar{\omega}_i = 0 \quad \text{on } U.$$

The calculus volume element induced by $\bar{\omega}$ on $U \subset M_t$ is the t dependent measure $\mu_{\bar{g}_t}$ if the orientation of U, as element of $\partial\Omega$, is the orientation induced by the ordering of the coordinates $x^1, \ldots, x^n$,

$$\mu_g := |\det g|^{\frac{1}{2}} dx^1 \ldots dx^n.$$

In case of an opposite orientation of $\partial\Omega$, $\bar{\omega}_\alpha$ induces a volume element of opposite sign. In particular on the boundaries S_t and S_0 we have, denoting by integrals over subsets of $\partial\Omega$ integrals of exterior forms on oriented manifolds and suppressing the mention of $\partial\Omega$ in calculus integrals

$$\int_{\partial\Omega \cap M_t} \mathcal{P}.\bar{\omega} = \int_{S_t} \mathcal{P}^0 N \mu_g,$$

$$\int_{\partial\Omega \cap M_0} \mathcal{P}.\bar{\omega} = -\int_{S_0} \mathcal{P}^0 N \mu_g.$$

We have[7] $\mathcal{P}^0 N = \mathcal{P}^\alpha n_\alpha$ if n is the past oriented normal to S_t. It is positive and equal to the density of energy $\varepsilon_{X,t}(u)$ if $\mathcal{P}$ is future oriented; that is if X is past oriented like n. We then have:

$$\varepsilon_{X,t}(u) := \mathcal{P}^\alpha n_\alpha \equiv \frac{1}{2} \mid \partial u \mid^2_{G_{X,n}} \equiv \frac{1}{2} G^{\alpha\beta}_{X,n} \partial_\alpha u \partial_\beta u.$$

The **energy** of u on S_t, relative to the timelike vector field X, is the integral:

$$\mathcal{E}_{X,S_t}(u) := \int_{S_t} \varepsilon_{X,t}(u) \mu_g. \tag{2.27}$$

We have proved the following theorem.

Theorem 2.10 ***(Energy equality)*** *Let u be a C^2 solution of the wave equation (2.1) on a* [8] *compact domain, Stokes regular, Ω of $M \times [0,t]$, of a sliced C^1*

[7] Remark that the contravariant zero components of a vector are the same in the natural frame and the Cauchy adapted frame.

[8] i.e. a domain with compact closure.

Lorentzian manifold, such that the boundary $\partial\Omega$ is the union of compact subsets S_t and S_0 of respectively M_t and M_0, and a lateral boundary L. Then

$$\int_{S_t} \varepsilon_{X,t}(u)\mu_{\bar{g}} + \int_L \mathcal{P}.\ell\mu_L - \int_{S_0} \varepsilon_{X,0}(u)\mu_{\bar{g}} = \int_0^t \int_{S_\tau} \left\{ fX.\partial u + \frac{1}{2}T.\mathcal{L}_X(g) \right\} N\mu_{\bar{g}} d\tau$$

We will deduce from this energy equality a fundamental energy inequality.

2.5 *Energy inequality in a compact causal domain*

Definition 2.11 *We say that a Stokes regular compact domain Ω of the strip $M \times [0,\tau]$ of the spacetime V is a* ***causal subset based on*** *$S_0 \subset M_0$ if its boundary is constituted of relatively compact domains S_τ of M_τ and S_0 of M_0 and a lateral boundary L which is spacelike or null and ingoing, i.e. every timelike line entering Ω is past directed.*

Fundamental compact causal subsets are intersections of pasts of subsets S_τ with the future, $t \geq 0$, of M_0. The general construction of such subsets requires a geometrical study of pasts and futures, which is done in Chapter 12. We are concerned in this appendix only with local properties, which rely on the following basic example.

Basic example. Let U be the domain of a coordinate chart of V with domain of local coordinates containing the product $B_\rho \times [0,\tau]$, with B_ρ a ball of R^n of centre 0 and radius ρ. In the compact domain $B_\rho \times [0,\tau]$ the lapse N is bounded above and below:

$$0 < A \leq N \leq B,$$

we take the shift $\beta = 0$ for simplicity. In B_ρ the induced metrics $\bar{g}_t$ are uniformly equivalent to the Euclidean metric:

$$a\Sigma(x^i)^2 \leq g_{ij}x^i x^j \leq b\Sigma(x^i)^2, \quad a > 0.$$

Take for S_τ a compact domain of M_τ with image a ball $B_\tau \subset R^n$, of centre 0 and radius $\rho_\tau < \rho$. The truncated cone

$$t - \tau = -\left(\frac{a}{B}\right)^{\frac{1}{2}} (r - \rho_\tau), \quad r^2 := \Sigma(x^i)^2$$

is the image of a submanifold $L \subset U$, intersecting M_0 as the boundary of a compact domain S_0 with image the ball $r \leq \rho_0 = \rho_\tau + (\frac{B}{a})^{\frac{1}{2}}\tau$, if τ is small enough to give $\rho_0 \leq \rho$, which limits the domain of the chart. The submanifold L is spacelike and ingoing. The compact domain $\bar{\Omega}$ with boundaries S_τ, S_0, and L is causal.

To bound the G norm of the stress energy tensor T in terms of the energy density we will use the following lemmas which are straightforward to prove:

Lemma 2.12 *Denote by $|\ \ |$ the norm of a covariant tensor in some positive definite quadratic form G, or of a contravariant tensor in the inverse quadratic form G^{-1}. Let u and v be two tensors; then*

$$|\, u \otimes v \,| = |\, u \,||\, v \,|, \quad |\, u.v \,| \leq |\, u \,||\, v \,|\,.$$

We denote by C numbers which depend only on the dimension of the manifold V.

Lemma 2.13 *The $G_{X,n}$norms of $\mathbf{g}$ and its contravariant associate $\mathbf{g}^{-1}$*

$$|\mathbf{g}|^2_{G(X,n)} \leq C\mathbf{g}(X,n)^2.$$
$$|\mathbf{g}|^2_{G^{-1}(X,n)} \leq C\mathbf{g}(X,X)^{-2},$$

denoting by C some number depending only on n.

We deduce from these lemmas a bound of the G norm of the stress energy tensor T in terms of the energy density, namely:

$$\mid T\mid_{G(X,n)} \leq \mid \partial u \otimes \partial u \mid_{G(X,n)} + \frac{1}{2}|\mathbf{g}| \mid \mathbf{g}\ (\partial u.\partial u) \mid_{G(X,n)}$$

hence

$$\mid T \mid_{G(X,n)} \leq C \mid \partial u \mid^2_{G(X,n)}$$

We denote by $\|f\|_\tau$the L^2 norm of the source f on (S_τ, g) and by $p(\tau)$ the supremum on S_τ of the $G_{X,n}$ norm of $\mathcal{L}_X\mathbf{g}$:

$$\| f \|_t = \left\{ \int_{S_t} \mid f \mid^2 \mu_g \right\}^{1/2} \tag{2.28}$$

$$p(t) \equiv Sup_{S_t} \mid \mathcal{L}_X\mathbf{g} \mid_{G(X,n)} \tag{2.29}$$

and we will prove the following fundamental theorem.

Theorem 2.14 ***(local energy inequality)*** *If $\mathbf{g}$ and X are C^1 while f is C^0, any C^2 solution of the wave equation with source f satisfies on a compact causal domain the fundamental energy inequality:*

$$\mathcal{E}_{X,S_t}(u)^{1/2} \leq C_0(t) \left\{ \mathcal{E}_{X,S_0}(u)^{1/2} + \frac{1}{2}\int_0^t C_1(\tau) \| f \|_\tau \, d\tau \right\} \tag{2.30}$$

with $C_0(t)$ a continuous and increasing function of t, equal to 1 for $t = 0$ given by:

$$C_0(t) = exp\left[\frac{1}{2}\int_0^t C_2(\tau)p(\tau)d\tau\right] \tag{2.31}$$

The functions C_1 and C_2 are continuous and bounded functions of τ given by:

$$C_1(\tau) = Sup_{S_\tau}(N \mid X \mid_{G_{X,n}}), \quad C_2(\tau) = Sup_{S_\tau}\left(\frac{3}{2}N \mid \mathbf{g} \mid_{G_{X,n}}\right). \tag{2.32}$$

Proof We know that $\mathcal{P}$ is timelike or null, hence $\mathcal{P}.\ell \geq 0$ on L if ℓ is timelike or null with the orientation opposite to $\mathcal{P}$, which is the case due to our hypothesis on ℓ. To estimate the $n+1$ integral we bound the absolute value of the quantity

to integrate through the $G_{X,n}$ norms of its various factors. We deduce from the integral formula the definition of $\varepsilon_{X,t}(u)$ and the Schwartz inequality that:

$$\mathcal{E}_{X,S_t}(u) \leq \mathcal{E}_{X,S_0}(u) + \int_0^t \{C_1(\tau) \parallel f \parallel_\tau \mathcal{E}_{X,S_\tau}(u)^{1/2} + C_2(\tau)p(\tau)\mathcal{E}_{X,S_\tau}(u)\}d\tau. \tag{2.33}$$

A positive and continuous solution $\mathcal{E}_{X,S_t}(u)$ of this integral inequality, which has positive coefficients, is bounded by a solution of the associated integral equality (and elementary calculus result called the **Gromwall lemma**); hence $\mathcal{E}_{X,S_t}(u)$ is bounded by a solution y of the differential equation

$$y'(t) = C_1(t) \parallel f \parallel_t y(t)^{1/2} + C_2(t)p(t)y(t),$$

if y takes an initial value $y(0) \leq \mathcal{E}_{S_0}(u)$. Elementary calculus gives for this solution the function y such that

$$y^{1/2}(t) = exp\left(\frac{1}{2}\int_0^t C_2(\tau)p(\tau)d\tau\right) \left\{y^{1/2}(0) + \frac{1}{2}\int_0^t \left[C_1(\tau) \parallel f \parallel_\tau exp\left(-\frac{1}{2}\int_0^\tau C_2(\theta)p(\theta)d\theta\right)\right] d\tau\right\}.$$

The estimate of the theorem follows from this value of the majorant y, since the exponential of a negative number, here $-\frac{1}{2}\int_0^\tau C_2(\theta)p(\theta)d\theta$, is less than or equal to 1. □

2.6 *Uniqueness theorem and causality*

Consider the Cauchy problem for the wave equation (2.1), that is a solution with given initial data:

$$u(0,.) = \varphi, \quad (\partial_0 u)(0,.) = \psi.$$

The difference of two solutions of this problem satisfies the homogeneous equation:

$$\Box_{\mathbf{g}}(u_1 - u_2) = 0$$

with zero initial data. The energy inequality implies, whatever the choice of X,

$$\mathcal{E}_{X,S_\tau}(u_1 - u_2) = 0 \quad \text{for} \quad \tau \in [0,t].$$

It implies therefore $\partial(u_1 - u_2) = 0$ on Ω, and also $u_1 - u_2 = 0$ since it is so on S_0. We have proven:

Theorem 2.15 *(**Uniqueness and causality**) Two C^2 solutions of the wave equation on a C^1 sliced Lorentzian manifold which take the same initial data $u(0,.) = \varphi$, $\partial_t u(0,.) = \psi$ on a compact subset S_0 of M_0 coincide on any compact causal domain of V based on S_0.*

This theorem will be extended later to less regular data. We will also prove the continuous dependence of the solution on initial data and on f.

2.7 *Case of a strip*

We denote by $V_T := M \times [0,T]$ a strip of V. We denote by $C_0^k(T)$ the restriction to V_T of C^k functions with compact support in V. A function in $C_0^k(T)$ has a support in $K \times [0,T]$, K some compact of M.

Let $u \in C_0^2(T)$ be a solution of the wave equation with support in $K \times [0,T]$. Let $\Omega = K' \times [0,T]$, K' a compact domain Stokes regular, strictly containing K. The field u and all its derivatives vanish outside of $K' \times [0,T]$, hence the only part of $\partial\Omega$ furnishing a contribution to the integral on $\partial\Omega$ in the Stokes formula applied to $\nabla_\alpha \mathcal{P}^\alpha$ will be contained in M_t and M_0, spacelike manifolds with timelike normals n.

The energy $\mathcal{E}_{X,M_t}$ relative to X of the field u on M_t is denoted $\mathcal{E}_{X,t}(u)$. It is given by

$$\mathcal{E}_{X,t}(u) \equiv \int_{M_t} G_{X,n}(\partial u, \partial u)\mu_g. \tag{2.34}$$

The integral equality obtained in the previous section now reads:

$$\mathcal{E}_{X,t}(u) = \mathcal{E}_{X,0}(u) + \int_0^t \int_{M_\tau} \left(fX.\partial u + \frac{1}{2} T.\mathcal{L}_X(\mathbf{g}) \right) N \mu_g d\tau. \tag{2.35}$$

In the case where X is a Killing vector field of $\mathbf{g}$, i.e. $\mathcal{L}_X \mathbf{g} = 0$, and when $f = 0$ (homogeneous wave equation) the equality above expresses the **conservation of energy** of the field u.

When the source f is not zero or X is not a Killing field of $\mathbf{g}$ we will deduce from the integral equality (2.35) an energy inequality modulo some additional hypothesis on the global in space properties of $\mathbf{g}$ and X which we now formulate.

Definition 2.16 *$X-$ energy uniformity hypothesis.*

1a. *N and $\mid \mathbf{g}(X,n) \mid = \gamma$ are uniformly bounded on each M_t (the two conditions coincide if $X = \partial/\partial t$). Then $\mid \mathbf{g}(X,X) \mid = \alpha^2$ is also uniformly bounded, since it is less than or equal to $\mathbf{g}(X,n)^2 = \gamma^2$.*

1b. *X is uniformly strictly timelike on each M_t.*
The hypotheses 1a and 1b mean that we postulate the existence of continuous and positive functions A_1, Γ_1, Γ_2 of $t \in [0,T]$ such that:

$$Sup_{M_t} N \leq A_1(t), \quad 0 < \Gamma_2(t) \leq Sup_{M_t} \alpha \leq Sup_{M_t} \gamma \leq \Gamma_1(t).$$

2. *$\mathcal{L}_X(\mathbf{g})$ is uniformly bounded in $G_{X,n}$ norm on each M_t, for $t \in [0,T]$. We set*

$$Sup_{M_t} \mid \mathcal{L}_X(\mathbf{g})|_{G_{X,n}} = p(t)$$

For the choice $X = n$ the above hypotheses, called simply the uniformity hypotheses, reduce to the following ones.

Definition 2.17 *Uniformity hypotheses.*

1. *N is uniformly bounded on each M_t.*
2. *The space gradient $\bar{\nabla} N$ and the extrinsic curvature K are uniformly bounded on each M_t.*

Theorem 2.18 *If $\mathbf{g}$ and X are C^1 and satisfy the X-energy uniformity hypothesis while $f \in C_0^0$, then the energy on M_t of a $C_0^2(T)$ solution of the wave equation satisfies the inequalities (2.35) for $0 \leq t \leq T$ with the functions C, C_1 and C_2 depending only on the uniformity bounds. In particular:*

$$C_1(t) \leq 2\Gamma_1(t), \qquad C_2(t) \leq A_1(t) n^{1/2} \Gamma_2^{-1/2}(t) \tag{2.36}$$

Proof Follows the same lines as in a previous section and uses the uniform bounds of the $G_{X,n}$ norms of the various involved quantities. □

Remark 2.19 *We will extend later the energy inequality to bigger spaces than C_0^2 for u.*

2.8 *Estimate of u*

The energy inequality gives only an estimate of ∂u.

- An estimate of u can be obtained from its initial data by using the formula which defines u as a function such that $u - u(.,0) \in C^1([0,T], H_0)$ when $\partial_t u \in C^0([0,T], H_0)$, with H_0 the space of square integrable functions in some given volume element on M:

$$u(.,t) - u(.,0) = \int_0^t \partial_t u(.,\tau) d\tau. \tag{2.37}$$

  This formula implies, by the Cauchy–Schwartz inequality, that:

$$\| u(.,t) - u(.,0) \| \leq t^{1/2} \left\{ \int_0^t \| \partial_t u(.,\tau) \|^2 \, \partial\tau \right\}^{1/2} \tag{2.38}$$

  with $||.||$the norm in H_0.
- In the case of the wave equation with positive mass m,

$$\Box_{\mathbf{g}} u - mu = f, \qquad m > 0, \tag{2.39}$$

  one obtains directly L^2 estimates of u on M_t by adding to the previous stress energy tensor the term $-\frac{1}{2}\mathbf{g} m u^2$ and replacing the density of energy by

$$\varepsilon_t(u) := | \partial u |^2_{G_{X,n}} + \frac{1}{2} | \mathbf{g}(X,n) | m u^2. \tag{2.40}$$

2.9 *Cauchy problem*

We consider the wave equation with mass m and source f on the s.L. manifold $(V, \mathbf{g})$, with Cauchy data on M_0:

$$\Box_{\mathbf{g}} u - mu = f,$$

$$u(.,0) = \varphi, \quad \partial_t u(.,0) = \psi.$$

2.9.1 *Uniqueness theorem*

Theorem 2.20 *If the metric* $\boldsymbol{g}$ *satisfies the uniformity hypothesis (2.17), the Cauchy problem with data* $\varphi \in C_0^2, \psi \in C_0^1$ *has at most one solution* $u \in C_0^2(T)$*, whatever be the interval* T*.*

Proof Theorem 2.15. □

2.9.2 *Existence theorem*

The local in time[9] existence theorem can be deduced from the Cauchy–Kovalevski theorem which applies to equations with analytic coefficients and analytic Cauchy data. One approximates non-analytic coefficients and data by analytic ones and uses energy inequalities to show the convergence of the analytic solutions to a solution of the original, non-analytic problem. It is a solution in the sense of distributions, having second derivatives only in this generalized sense, unless coefficients and data are more regular and one has proven higher order energy estimates.

The energy inequality leads directly to an existence theorem on $V_T = M \times [0 \times T]$, for any T, of a solution through Hilbert space methods[10], again in the generalized sense. We give the proof below, postponing the proof of higher order energy estimates to the next section.

In order to be able to readily extend the results to general linear systems, and to use them for non-linear equations, we introduce a given, fixed, C^∞, properly Riemannian metric e on M. We use some definitions and fundamental properties given in Appendix I.

Definition 2.21

1. *The space* H_s *on* (M, e) *is the space of functions, or tensor fields of some given type, on* M *with square integrable generalized derivatives of order* $\leq s$ *in the metric* e*. It is a Hilbert space with norm*

$$\| f \|_{H_s} := \left\{ \int_M \sum_{0 \leq k \leq s} | \partial^k f |^2 \, \mu_e \right\}^{1/2} \tag{2.41}$$

[9] Its global in time extension for linear equations can also be obtained (Leray, 1953).

[10] Initiated by K. O. Friedrichs in 1954 for the case of first-order symmetric hyperbolic systems.

where $|\ |$, ∂, and μ_e denote respectively the pointwise norm, covariant derivative, and volume element in the metric e.

2. *The Sobolev (Hilbert) spaces $\overset{0}{H}_s \subset H_s$ on M are closures of the space $\mathcal{D}$ of C^∞ functions, or tensor fields of some given type, with compact support in M with respect to the norms (2.41).*

Proposition 2.22[11]

1. *The space $\overset{0}{H}_0$ is identical with the space $H_0 \equiv L^2(M, e)$ of square integrable functions on M in the volume element of e.*

2. *The space $\overset{0}{H}_1$ coincides with the space H_1 when (M, e) has a non-zero injectivity radius and hence is complete. It is true in particular if M is a compact manifold (without boundary). On the other hand, if (M, e) can be identified with a relatively compact subset, with differentiable boundary ∂M, of a smooth Riemannian manifold (M', e') then functions in $\overset{0}{H}_1$ have a zero trace on ∂M: they constitue only a subset of H_1.*

Definition 2.23

1. *We denote by $E_s(T)$ the following Banach spaces:*

$$E_s(T) \equiv \cap_{0\leq k\leq s} C^{s-k}([0,T], H_k)$$

2. *We denote by $\overset{0}{E}_s(T)$ the Banach spaces*

$$\overset{0}{E}_s(T) \equiv \cap_{0\leq k\leq s} C^{s-k}([0,T], \overset{0}{H}_k).$$

They are the completion of $C_0^s(T)$ in the $E_s(T)$ norm:

$$\| u \|_{E_s(T)} = Sup_{t\in[0,T]} Sup_{0\leq k\leq s} \| (\partial_t^k u)(.,t) \|_{H_{s-k}}, \quad \partial_t \equiv \frac{\partial}{\partial t}. \tag{2.42}$$

Define the metric $\mathbf{e}$ on $V = M \times R$ to be:

$$\mathbf{e} = dt^2 + e,$$

and denote by $|\ |$ and by $\mathbf{D}$ the pointwise $\mathbf{e}$ norm and $\mathbf{e}$ covariant derivative of a spacetime tensor, and by $\mathbf{D}^k$ the covariant derivative of order k. Equivalent definitions of the $E_s(T)$ norm are:

$$\| u \|_{E_s(T)} = Sup_{t\in[0,T]} \left\{ \sum_{0\leq k\leq s} \int_{M_t} |\ \mathbf{D}^k u|^2 \mu_e \right\}^{1/2} \tag{2.43}$$

[11] See Appendix I.

$$\| u \|_{E_s(T)} = \sum_{0\leq k\leq s} \| \mathbf{D}^k u \|_{E_0(T)},$$

$$\| u \|_{E_s(T)} = Sup_{0\leq k\leq s} \| \mathbf{D}^k u \|_{E_0(T)} .$$

In order to relate the energy previously defined to our H_1 and E_1 norms we make a hypothesis on the spacetime metric **g**.

Definition 2.24 *Fundamental Regularity Hypothesis*[12].

1. *The metrics g_t, $t \in [0,T]$, are equivalent to e, i.e. there exist continuous (strictly) positive functions $B_1(t), B_2(t)$ such that for all $t \in [0,T]$ and each tangent vector ξ to M it holds on M*

$$B_2(t)e(\xi,\xi) \leq g_t(\xi,\xi) \leq B_1(t)e(\xi,\xi) \tag{2.44}$$

The lapse N is such that there exist continuous (strictly) positive functions $A_1(t), A_2(t)$ on $[0,T]$ such that on each M_t it holds:

$$A_1(t) \geq N \geq A_2(t). \tag{2.45}$$

Since the functions B_1, B_2, A_1, A_2 are continuous they are bounded above and below by a positive number on the compact interval $[0,T]$.

*The shift β is uniformly bounded in e-norm on each M_t by a number $b(t)$. Such a manifold is said to be **regularly sliced**.*

2. *The spacetime metric **g** is C^1 on V*[13]*. The derivative* $\mathbf{Dg}$*, i.e. the derivatives $\partial_t \boldsymbol{g}$ and $D\mathbf{g}$, are uniformly bounded in* $\mathbf{e}$*-norm on V_T.*

To define the energy of u at time t we choose the timelike vector X to be the past-oriented unit normal n. The uniformity hypothesis (2.17) is a consequence of the Fundamental Regularity Hypothesis. Moreover:

Lemma 2.25

1. *Under the regularity hypothesis 1 of Definition 2.24, the quadratic form $G_{n,n}$ is uniformly equivalent to $\mid \xi_t \mid^2 + \mid \bar{\xi} \mid^2$, that is to the contravariant associate of **e**.*

2. *If the metric **g** is such that*[14]*, on M_0,*

$$N(.,0) = 1, \quad \beta(.,0) = 0 \tag{2.46}$$

[12] In the case of the wave equation the hypothesis on the bound of $\mathbf{Dg}$ can be replaced by a hypothesis on the bound on $\mathcal{L}_X \mathbf{g}$.

[13] Part 1 of our hypothesis implies $\mathbf{g} \in C_b^0(T)$. Part 2 could be replaced by $\mathbf{Dg} \in L^\infty(V)_T$

[14] It is always possible to satisfy these conditions by a change of trivialization of $V = M \times R$ in a neighbourhood of M_0.

and we choose

$$e = g(.,0) \tag{2.47}$$

then the positive quadratic form $G_{n,n}$ reduces on M_0 to

$$G_{n,n}(.,0) = \mathbf{e}. \tag{2.48}$$

Proof We have, with ξ_t the t-component of ξ in the natural coframe:

$$G_{n,n}(\xi,\xi) \equiv N^{-2}(\xi_0)^2 + g^{ij}\xi_i\xi_j \equiv N^{-2}(\xi_t - \beta^i\xi_i)^2 + g^{ij}\xi_i\xi_j$$

from which follows conclusion 2.

It holds that

$$|\xi_t - \beta^i\xi_i| \leq |\xi_t| + b|\xi|.$$

Therefore, using hypothesis 2.24,

$$G_{n,n}(\xi,\xi) \leq 2(A_2)^{-2}(\xi_t)^2 + (2b^2 + B_1)|\bar{\xi}|^2$$

Also

$$\frac{1}{2}\{|\ \xi_t\ |^2 + |\ \bar{\xi}\ |^2\} \leq \xi_0^2 + b^2|\bar{\xi}|^2 + |\bar{\xi}|^2 \leq A_1^{-2}N^{-2}\xi_0^2 + (b^2+1)B_2^{-1}g^{ij}\xi_i\xi_j.$$

Conclusion 1 follows immediately. □

We now estimate the energy on M_t in terms of the given norms. To simplify the writing we consider the case of the wave equation with mass $m > 0$. The case of zero mass can be treated along the same lines by using the inequality of Section 2.8 or introducing variants of the functional spaces E_s, less restrictive for u on non-compact space manifolds.

Lemma 2.26 *Under hypothesis 1 of Definition 2.24 the energy of u on M_t, with $X = n$, is uniformly equivalent for $t \in [0,T]$ to the square of the sum of the H_1 norm of $u(.,t)$ and the H_0 norm of $\partial_t u(.,t)$, i.e. there exist positive and continuous functions of t, $C_1(t)$ and $C_2(t)$ such that*

$$C_1(t)y(t) \leq \mathcal{E}_t^{1/2}(u) \leq C_2(t)y(t),$$

where

$$y(t) := \frac{1}{2}\{\|\ u(.,t)\ \|_{H_1}^2 + \|\ \partial_t u(.,t)\ \|_{H_0}^2)\}^{\frac{1}{2}}.$$

Proof Lemma 2.25 and the fact that $\mathbf{g}(n,n) = -1$, together with the uniform equivalence of the volume elements of g and e. □

Under the Fundamental Regularity Hypothesis 2.24, the operator $\Box_{\mathbf{g}} - m$, which maps $C_0^2(T)$ into $C_0^0(T)$, extends by continuity to an operator from $\overset{0}{E}_2(T)$ into $\overset{0}{E}_0(T)$.

The space $\overset{0}{E}_0(T)$ is included into $L^2(T)$, Hilbert space of square integrable functions on V_T.

We denote by L the linear operator from a Banach into a Hilbert space defined as follows:

$$L : \overset{0}{E}_2(T) \to L^2(T) \times \overset{0}{H}_1 \times H_0) \quad \text{by} \quad u \mapsto (\mathbf{g}.\nabla^2 u - mu), u(.,0), (\partial_t u)(.,0)).$$

Let us denote by $\mathcal{R} \subset L^2(T) \times \overset{0}{E}_1 \times H_0$ the range of L.

Lemma 2.27 *The inverse of the operator L is a continuous map from $\mathcal{R}$ with the topology $L^2(T) \times \overset{0}{H}_1 \times H_0$ into the Banach space $\overset{0}{E}_1$.*

Proof By its definition L^{-1} maps $\mathcal{R}$ onto $\overset{0}{E}_2$, hence into $\overset{0}{E}_1$ which contains $\overset{0}{E}_2$ as a subset. We have to show that L^{-1} is continuous for the topology given to $\mathcal{R}$ and the topology E_1 given to its range $\overset{0}{E}_2$. This continuity is a consequence of the energy inequality and Theorem 2.14, which imply:

$$\| u \|_{E_1(T)} \leq C(T)\{y(0) + \| f \|_{L^2(T)}\}$$

with:

$$y(0) = \{\| u(.,0) \|^2_{H_1} + \| \partial_t u(.,0) \|^2_{H_0}\}$$

It results from the lemma that L^{-1} extends to a continuous map $\bar{L}^{-1}$ from the closure $\bar{\mathcal{R}}$ of $\mathcal{R}$ (in its topology) into $\overset{0}{E}_1$. □

Lemma 2.28 *We shall prove the following:*

1. *If $u = \bar{L}^{-1}(f, \varphi, \psi)$, with $(f, \varphi, \psi) \in \bar{\mathcal{R}}$, then $\Box_{\mathbf{g}} u - mu = f$ in the generalized sense, and $u(.,0) = \varphi$, $(\partial_t u)(.,0) = \psi$.*

2. *The closure $\bar{\mathcal{R}}$ of $\mathcal{R}$ in $L^2(T) \times \overset{0}{H}_1 \times H_0$ is the whole of this space.*

Proof

1. If $(f, \varphi, \psi) \in \bar{\mathcal{R}}$ there exists a sequence $u_n \in E_2(T)$ such that

$$\Box_{\mathbf{g}} u_n - mu_n = f_n, u_n(.,0) = \varphi_n (\partial_t u)(.,0) = \psi_n$$

(f_n, φ_n, ψ_n) converges to (f, φ, ψ) in $L^2(T) \times \overset{0}{H}_1 \times H_0$.

We deduce from the energy inequality that the sequence u_n converges in E_1 to a function u, with trace φ on M_0, while $\partial_t u$ has trace ψ. Since u converges a fortiori in the sense of distributions it satisfies $\Box_{\mathbf{g}} u - mu = f$ in this sense.

2. We will have $\bar{\mathcal{R}} = \mathcal{H} \equiv L^2(T) \times \overset{0}{H}_1 \times H_0$ if its orthogonal complement in the scalar product of this Hilbert space reduces to zero. An equivalent condition since $\mathcal{R}$ is dense in $\bar{\mathcal{R}}$ is that any given $v = (h, \Phi, \Psi) \in \mathcal{H}$ is zero if it satisfies

the condition

$$(v,w)_{\mathcal{H}} \equiv (h,f)_{L^2(T)} + (\Phi,\varphi)_{H_1} + (\Psi,\psi)_{H_0} = 0, \quad \forall w = (f,\varphi,\psi) \in \mathcal{R}$$

that is, by the definition of $\mathcal{R}$

$$\forall u \in \overset{0}{E}_2, \quad f = \Box_{\mathbf{g}} u - mu, \quad u(.,0) = \varphi, (\partial_t u)(.,0) = \psi$$

Take for $w = (f,\phi,\psi)$ an element of $\mathcal{R}$ such that $u \in \overset{0}{E}{}^{+}_{2}(T)$, subspace of functions of $\overset{0}{E}_2(T)$ which vanish together with their first derivatives on M_0. The scalar product (v,w) reduces then to:

$$(v,w)_{\mathcal{H}} = (h, \Box_{\mathbf{g}} u - mu)_{L^2(T)}$$

Suppose that $h \in \overset{0}{E}{}^{-}_{2}(T)$ subspace of $\overset{0}{E}_2(T)$ of functions which vanish together with their first derivatives on M_T, then, using integration by parts, Stokes formula and the conditions satisfied by u and h respectively on M_0 and M_T we find

$$(h, \Box_{\mathbf{g}} u - mu)_{L^2(T)} = (u, \Box_{\mathbf{g}} h - mh)_{L^2(T)}$$

We see that if $h \in \overset{0}{E}{}^{-}_{2}(T)$ the equality $(v,w)_{\mathcal{H}} = 0$ implies $(u, \Box_{\mathbf{g}} h - mh)_{L^2(T)} = 0$, $\forall u \in \overset{0}{E}{}^{+}_{2}(T)$, hence also for all $u \in L^2(T)$ since $\overset{0}{E}{}^{+}_{2}(T)$ is dense in that space, therefore

$$\Box_{\mathbf{g}} h - mh = 0.$$

This equation together with the hypothesis $h \in \overset{0}{E}{}^{-}_{2}(T)$ implies, by the uniqueness theorem, that $h = 0$ on V_T.

The previous proof together with the density of $\overset{0}{E}{}^{-}_{2}(T)$ in $L^2(T)$ and continuity of the scalar product shows that $(v,w)_{\mathcal{H}} = 0$, $\forall w \mathrm{w} \in \mathcal{R}$ implies $v = (0, \Phi, \Psi)$. Therefore $(v,w)_{\mathcal{H}} = 0$ reduces to

$$(\Phi,\varphi)_{H_1} + (\Psi,\psi)_{H_0} = 0, \qquad \forall (\varphi,\psi) \in \overset{0}{H}_2 \times \overset{0}{H}_1$$

since the traces $\varphi = u(0)$, $\psi = (\partial_t u)(.,0)$, of an arbitrary u in $\overset{0}{E}_2(T)$ are arbitrary elements of $\overset{0}{H}_2 \times \overset{0}{H}_1$. The equation above implies that any given pair $(\Phi, \breve{\Psi}) \in \overset{0}{H}_1 \times \overset{0}{H}_0$ is zero, because $\overset{0}{H}_s$ is dense in $\overset{0}{H}_{s-1}$.

The surjectivity of $\bar{L}^{-1}$ from $\overset{0}{E}_1(T)$ onto $L^2(T) \times H_1 \times H_0$, has been proven. □

We have proven the following theorem.

Theorem 2.29 *The Cauchy problem with $\overset{0}{H}_1 \times H_0$ Cauchy data for the wave equation*

$$\Box_{\mathbf{g}} u - mu = f, \quad m \geq 0$$

on $(M \times R, \mathbf{g})$, with $\boldsymbol{g}$ satisfying the Fundamental Regularity Hypothesis 2.24 and $f \in L^2(T)$, has a global solution u with $u \in E_1(T)$ for any finite T.

Remark 2.30 *By its construction the solution obtained in $\overset{0}{E}_1(T)$ for $H_1 \times H_0$ initial data and $f \in L^2(T)$ satisfies the energy inequality, but we have not proved that all solutions in $\overset{0}{E}_1(T)$ satisfy this inequality. However, it is possible to prove a uniqueness theorem in $\overset{0}{E}_1(T)$, for $\overset{0}{H}_1 \times H_0$ initial data*[15].

The existence and uniqueness theorems extend, with the same type of proof, to the case of a zero mass m, with Cauchy data in $\overset{0}{H}_1 \times H_0$, or in other spaces where $D\varphi$ and ψ are in H_0, while $\varphi - \varphi_0 \in H_0$, with φ_0 some given function on M_0.

2.10 *Generalizations*

The existence theorem that we have proved on a strip restricts, if M is not a compact manifold, the behaviour in space of the Cauchy data and of the solution.

- If (M, g_t) is not complete, then (M, e) cannot be chosen complete either since it is supposed to be equivalent to (M, g_t). The hypothesis $\varphi \in \overset{0}{H}_1$ implies that if M admits a boundary ∂M in a larger manifold, then φ vanishes on ∂M. Then the solution $u(t, .)$ vanishes on ∂M: the theorem solves an initial–boundary value problem.
- If (M, e) is complete, for instance Euclidean at infinity, then $\overset{0}{H}_1$ coincides with H_1, but square integrability of the field may be too strong a requirement to be satisfactory for physics. On the Euclidean space R^3, physical fields decrease typically like r^{-1} at infinity, they are not square integrable[16]. This difficulty can be remedied, in the case of the wave equation with $m = 0$, by taking Cauchy data such that only $\varphi - \varphi_0 \in H_0$, with φ_0 a given function.
- We will in the next section prove existence in other functional spaces, in particular in locally defined spaces, using compact causal domains, but we will work directly for more general linear systems. The obtained solutions are global in time only under uniformity hypotheses.

We will also prove the existence of more regular solutions, for more regular coefficients and Cauchy data and continuous dependence on the data.

[15] Cagnac, F. and Choquet-Bruhat, Y. (1984), *J. Math. Pures Appl.*, **63**, 377–90.
[16] Though, for $n > 4$, continuous fields decaying like r^{2-n} are square integrable on R^n.

3 Quasidiagonal linear systems

3.1 *Definitions*

To define general differential operators on a manifold $V \equiv M \times R$, and also to have a proper norm for tensor fields we use on V the Riemannian metric:

$$\mathbf{e} \equiv dt^2 + e \tag{3.1}$$

with e a smooth Riemannian metric on M. We denote by $\mathbf{D} = (\partial/\partial t, D)$ the covariant derivative in the metric $\mathbf{e}$ on V, the covariant derivative in e on M being denoted by D. Boldface is not used when Greek or Latin indices make clear the distinction between objects on V or M.

We consider a general linear quasidiagonal second-order system on the s.L. (sliced Lorentzian) manifold (V,$\mathbf{g}$), which we write for later convenience in a form which makes a term with constant mass m appear:

$$g^{\alpha\beta} D^2_{\alpha\beta} u + a^\alpha D_\alpha u + (b - m)u = f \tag{3.2}$$

where u and f are sections of a vector bundle $\mathcal{V}$ over V (for example u is a set of scalar functions or a tensor field), a is a vector-valued linear mapping from sections of $T^* \otimes \mathcal{V}$ into sections of $\mathcal{V}$, b is a linear map from sections of $\mathcal{V}$ into themselves, and m is a positive number.

3.2 *Stress energy tensor*

We denote by an underlined dot the pointwise scalar product relative to $\mathbf{e}$ and we define, by analogy with Definition 2.7, the stress energy tensor of u to be the symmetric 2-tensor:

$$T_{\alpha\beta} = D_\alpha u \underline{.} D_\beta u - \frac{1}{2} g_{\alpha\beta} (D_\lambda u \underline{.} D^\lambda u + m u \underline{.} u). \tag{3.3}$$

We consider as before a past-oriented timelike vector X. The energy momentum vector is (indices raised with $\mathbf{g}$)

$$\mathcal{P}^\alpha \equiv T^{\alpha\beta} X_\beta. \tag{3.4}$$

Its $\mathbf{e}$-divergence $\mathcal{P}$ is:

$$D_\alpha \mathcal{P}^\alpha \equiv D_\alpha (T^{\alpha\beta} X_\beta) \equiv X_\beta D_\alpha T^{\alpha\beta} + \frac{1}{2} T^{\alpha\beta} (D_a X_\beta + D_\beta X_\alpha) \tag{3.5}$$

with now

$$D_\alpha T^{\alpha\beta} \equiv D^\beta u \underline{.} (g^{\alpha\lambda} D_\alpha D_\lambda u - mu) + F^\beta, \tag{3.6}$$

where a straightforward computation gives

$$F^\beta \equiv D^\alpha u \underline{.} g^{\beta\lambda} (D_{\alpha\lambda} u - D_{\lambda\alpha} u) + D_\lambda u \underline{.} (D_\alpha g^{\alpha\lambda} D^\beta u + D_\alpha g^{\beta\lambda} D^\alpha u) \tag{3.7}$$
$$- \frac{1}{2} D_\alpha (g^{\alpha\beta} g^{\lambda\mu}) D_\lambda u . D_\mu u - \frac{1}{2} (D_\alpha g^{\alpha\beta}) m |u|^2.$$

If u is a scalar, or a set of scalar functions, the term $D_\alpha D_\lambda u - D_\lambda D_\alpha u$ is zero; if u is a tensor field this term is a linear map acting on u determined by the Riemann tensor of $\mathbf{e}$.

3.3 *Energy inequality in a compact causal domain*

We suppose that $\mathbf{g}$ satisfies the Fundamental Regularity Hypothesis 2.24, we take $X \in C^1$, then $\mathcal{P}^\alpha \equiv T^{\alpha\beta}X_\beta \in C^1$ if $u \in C^2$. We suppose $a, b \in C^0$.

Let Ω be a compact causal domain of $(V,\mathbf{g})$, Stokes regular, with boundary $\partial\Omega$ composed of $S_t \subset M_t$, $t > 0$, $S_0 \subset M_0$, and a lateral boundary supposed to be spacelike, or null and ingoing. We obtain, by analogy with what we did for the wave equation, the first energy local inequality for the solutions $u \in C^2(\Omega)$ of the given general linear system by integrating in Ω the divergence identity (3.5) with respect to the volume element

$$\mu_{\mathbf{e}} = dt\mu_e. \tag{3.8}$$

We find that[17]

$$\mathcal{E}_{X,S_t}(u) \leq \mathcal{E}_{X,S_0}(u) + \int_0^t \int_{S_\tau} \mid (f - a^\alpha D_\alpha u - bu)_{\cdot}(X.Du) + X.F + \frac{1}{2}T.\mathcal{L}_X\mathbf{e} \mid \mu(e)d\tau, \tag{3.9}$$

with

$$\mathcal{E}_{X,S_t}(u) := \int_{S_t} \varepsilon_{X,t}(u)\mu_e$$

and, ν being now the past-oriented unit normal to M_t in the metric $\mathbf{e}$, which has components $\nu_0 = \nu_t = 1, \quad \nu_i = 0$,

$$\varepsilon_{X,t}(u) \equiv \mathcal{P}^\alpha\nu_\alpha \equiv \mid \mathbf{D}u \mid^2_{G_{X,\nu}} + \frac{m}{2}|g(X,\nu)| \mid u \mid^2. \tag{3.10}$$

Theorem 3.1 *(Local uniqueness and causality) Under the given hypothesis on* $\boldsymbol{g}$*,* a*, and* b*, a* C^2 *solution of the Cauchy problem for our quasidiagonal linear system in a compact causal domain* Ω *depends only and continuously on the Cauchy data on its basis* S_0*.*

Proof Choose $X = \nu$. Suppose that $\mathbf{g}$ satisfies the Fundamental Regularity Hypothesis 2.24 on $\Omega \subset V_T$; then:

1. The quadratic form $G_{\nu,\nu}$ is in the Cauchy adapted frame (see 2.13)

$$G^{\alpha\beta}_{\nu,\nu}\xi_\alpha\xi_\beta \equiv N^{-2}\{N^{-2}\xi_0^2 + g^{ij}\xi_i\xi_j\}. \tag{3.11}$$

It is uniformly equivalent to the contravariant associate of $\mathbf{e}$ (see Lemma 2.25). On the other hand, due to the expression of $\mathbf{e}$ it holds that $\mathbf{D}_t u = \partial_t u$ and $\mathbf{D}_i u = D_i u$. Therefore, if $m > 0$, then $\varepsilon_{\nu,t}(u)$ is uniformly equivalent to

$$|\mathbf{D}u|^2 + |u|^2 = \mid \partial_t u \mid^2 + \mid Du \mid^2 + |u|^2$$

[17] Recall that a non-underlined dot denotes a contraction in the spacetime metric $\mathbf{g}$.

where $|.|$ denotes a norm in the metric $\mathbf{e}$. Therefore the function $\mathcal{E}_{\nu,S_t}(u)$ is uniformly equivalent for $t \in [0,T]$ to a function $y(t)$ given by

$$y(t) = \frac{1}{2}\int_{S_t} \{|\mathbf{D}u\,|^2 + |\,u\,|^2\}\mu_e. \tag{3.12}$$

2. Under the hypothesis of Lemma 2.25, it holds that

$$y(0) = \mathcal{E}_{\nu,S_0}(u). \tag{3.13}$$

With the choice $X = \nu$ the inequality (3.9) implies

$$y(t) \leq C_0(t)y(0) + \int_0^t \int_{S_\tau} |\,(f - a^\alpha D_\alpha u - bu)\underline{.}(\partial_t u) + F_t\,|\,\mu(e)d\tau$$

Therefore, under the Fundamental Regularity Hypothesis and the hypothesis made on a and b,

$$y(t) \leq C_0(t)y(0) + \int_0^t \|\,f\,\|_\tau\, y(\tau)^{\frac{1}{2}} + C(\tau)y(\tau)d\tau \tag{3.14}$$

with

$$\|\,f\,\|_\tau \equiv \left\{\int_{S_\tau} |\,f\,|^2\,\mu_e\right\}^{1/2}$$

from which we deduce by arguments similar to those of the proof of Theorem 2.14 an inequality of the form:

$$y(t)^{1/2} \leq C_0(t)\left\{y(0)^{1/2} + C(t)\int_0^t \|\,f\,\|_\tau\,d\tau\right\} \tag{3.15}$$

with C_0 and C positive and continuous function of t, depending on the bounds of the pointwise $\mathbf{e}$-norm of $\mathbf{g}$, $\mathbf{Dg}$, a, and b in the compact domain Ω. This inequality satisfied by the function y proves uniqueness, causality and continuous dependence on data for a C^2 solution of our Cauchy problem. □

3.4 *Case of a strip V_T*

We will prove the following theorem.

Theorem 3.2 *Suppose that on V_T the spacetime metric $\boldsymbol{g}$ satisfies the Fundamental Regularity Hypothesis 2.24, and that the coefficients a and b are continuous and uniformly bounded in $\boldsymbol{e}$ norm. Then a solution $u \in \overset{0}{E}_2(T)$ satisfies the following inequality, where we denote by $C_0(T)$, $C(T)$ various numbers depending only on the norms appearing in the hypothesis:*

$$\|\,u\,\|_{E_1(T)} \leq C_0(T)\{\|\,\varphi\,\|_{H_1} + \|\,\psi\,\|_{H_0}\} + C(T)\int_0^T \|\,f(.,t)\,\|_{H_0}\,dt \tag{3.16}$$

with

$$u(.,0) = \varphi, \qquad \partial_t u(.,0) = \psi. \tag{3.17}$$

The number $C_0(T)$ has the additional property to take the form

$$C_0(T) = 1 + TC(T) \tag{3.18}$$

under the hypothesis of Lemma 2.25 on the choice of e and the initial splitting of $(V, \mathbf{g})$.

Proof Suppose $u \in C_0^2(V_T)$. Set:

$$y(t) \equiv \frac{1}{2}\int_{M_t} \{\mid \mathbf{D}u \mid^2 + \mid u \mid^2\}\mu_e \equiv \frac{1}{2}\{||\partial_t u(.,t)||^2 + ||Du(.,t)||^2 + ||u(.,t)||^2\}$$

Under the hypothesis made on $\mathbf{g}$ the function $y(t)$ is uniformly equivalent to $\mathcal{E}_{M_t}(u)$. The function $y(0) \equiv \frac{1}{2}\{\| \varphi \|_{H_1} + \| \psi \|_{H_0}\}$ is equal to $\mathcal{E}_{M_0}(u)$ under the hypothesis (2.41). Since $\mathbf{Dg}$ is continuous and bounded the uniform equivalence coefficients of $y(t)$ and $\mathcal{E}_{M_t}(u)$ are of the form $1 + TC(T)$. The inequalities (3.14), (3.15) hold with S_τ replaced by M_τ, from which follows the announced inequality for $C_0^2(T)$ solutions of the linear system.

By our Definition 2.23 of the Banach space $\overset{0}{E}_2(T)$ the space $C_0^2(T)$ is dense in $\overset{0}{E}_2(T)$ On the other hand, with the hypothesis made on $\mathbf{g}$, a, and b we have that $f = Lu$ converges in[18] $E_0(T)$ when u converges in $\overset{0}{E}_2(T)$. Therefore the inequality holds for a solution $u \in \overset{0}{E}_2(T)$. □

Remark 3.3

1. *We have the following bounds*

$$\int_0^T \| f(.,t) \|_{H_0}\, dt \leq T \| f \|_{E_0} \tag{3.19}$$

also, by the Cauchy–Schwartz inequality:

$$\int_0^T \| f(.,t) \|_{H_0}\, dt \leq T^{1/2} \| f \|_{L^2(T)}\ .$$

2. *An inequality involving, on each M_t, only the L^2 norm of $\boldsymbol{D}u$, and not of u can be obtained if $b - m = 0$. Inequalities with L^p norms of u, with $p > 2$ can be obtained if $m = 0$ and b belongs to appropriate L^p spaces*[19]*. These extensions are important for physical problems where the rate of decay of u in space does not permit its belonging to an L^2 space. We will return to them for more regular solutions (see Section 3.6).*

[18] Recall that it always hold that $H_0 = \overset{0}{H}_0$.
[19] See Cagnac, F. and Choquet-Bruhat, Y. (1984) *lo cit.*

3.5 *Existence, uniqueness, causality and continuity*

3.5.1 *Uniqueness and causality*

The energy inequality (3.15) shows that two solutions with the same Cauchy data on the basis S_0 of a compact causal domain coincide in this domain, since their difference satisfies a homogeneous system ($f = 0$) with zero Cauchy data.

3.5.2 *Local and global existence*

The existence theorem 2.29 in $\overset{0}{E}_1(T)$ proved for the wave equation, for arbitrary T, extends to the systems considered in this section, using now the uniqueness theorem for the adjoint operator, which is an operator of the same type.

Suppose now that initial data are given on the basis S_0 of a compact causal domain $\Omega \subset (V, \mathbf{g})$. We extend them smoothly to functions in $H_1 \times H_0$ on M_0. The system admits a solution with these Cauchy data, which depend only on Ω, by the uniqueness theorem, on their restriction to S_0: this gives an existence theorem in the compact causal domain.

We have seen in Chapter 12 that every globally hyperbolic[20] manifold is a sliced Lorentzian manifold. This property and previous results give a global existence, uniqueness and continuous dependence on data result on such a manifold.

3.6 *Higher order estimates*

To solve non-linear problems it is necessary to prove the existence of solutions of linear equations with greater regularity than that obtained from the fundamental E_1 estimate. We will establish in this section E_s estimates for the general linear system of the previous subsection

$$Lu \equiv g^{\alpha\beta} D^2_{\alpha\beta} u + a^\alpha D_\alpha u + (b - m)u = f. \tag{3.20}$$

on an s.L. manifold $(V, \mathbf{g})$, $V \equiv M \times R$, endowed with a given Riemannian metric

$$\mathbf{e} = dt^2 + e, \tag{3.21}$$

where e is a smooth properly Riemannian metric on M and $\mathbf{D}$ the covariant derivative in $\mathbf{e}$.

3.6.1 *Functional spaces*

Definition 3.4 *We say that (M, e) is **Sobolev regular** if the spaces H_s satisfy the usual embedding and multiplication properties*[21] *which we state in the case of the Hilbert spaces H_s. The numbers $C(M, e)$ of the following inequalities, called Sobolev constants, depend only on M and e.*

[20] We have defined, following Leray, globally hyperbolic manifolds (Chapter 12) as Lorentzian manifolds where the set of timelike or null paths joining two points is compact in the set of paths.

[21] See Section 2.2 in Appendix I.

1. *The space H_s is continuously imbedded into $\bar{C}^0$ (continuous and bounded functions) if $s > \frac{n}{2}$,*

$$||u||_{H_s} \leq C(M,e) Sup_M |u|.$$

2. *The spaces H_s admit the continuous multiplication property*

$$H_{s_1} \times H_{s_2} \to H_s, \quad by \quad (u,v) \mapsto u \otimes v,$$

$$||u \otimes v||_s \leq C(M,e)||u||_{H_{s_1}}||u||_{H_{s_2}} \quad if \quad s_1 + s_2 > s + \frac{n}{2}, \quad s_1, s_2 \geq s$$

In particular H_s is an algebra if $s > \frac{n}{2}$.

Note that the multiplication property and the definition of $\overset{0}{H}_s$ imply:

$$\overset{0}{H}_{s_1} \times H_{s_2} \to \overset{0}{H}_s, \qquad \text{if} \quad s_1 + s_2 > s + \frac{n}{2}, \quad s_1, s_2 \geq s$$

We define for tensor fields on V spaces corresponding to the spaces $E_s(T)$ of functions of the previous section. The restriction to M_t of a p-tensor u on V decomposes, by projections on the tangent space to M_t and on the normal n, into a set of tensors on M, of orders $p, p-1, \ldots, 0$, depending on t: such decompositions are been given for the metric $\mathbf{g}$ and the Riemann tensor in Chapter 6. The restriction $u(.,t)$ is said to belong to H_s [respectively $\overset{0}{H}_s$] if each of the tensors of this decomposition is in H_s [respectively $\overset{0}{H}_s$] on (M,e). Recall, on the other hand, that due to the choice of $\mathbf{e}$ it holds that, in the $\mathbf{e}$ adapted frame:

$$D_0 u = D_t u = \partial_t u, \quad \mathbf{D}_i u = D_i u,$$

and $\mathbf{D}_0$ and $\mathbf{D}_i$ commute.

Definition 3.5

1. *We denote by $E_s(T)$ a Banach space of tensor fields on the strip V_T, each one equivalent to a set of t-dependent tensor fields on M, defined as follows:*

$$E_s(T) \equiv \cap_{0 \leq k \leq s} C^k([0,T], H_{s-k}) \tag{3.22}$$

The norm is:

$$\| u \|_{E_s(T)} \equiv Sup_{0 \leq t \leq T} \Sigma_{0 \leq k \leq s} \| \partial_t^k u(.,t) \|_{H_{s-k}}$$

The restriction to M_t of a p-tensor u on V decomposes, by projections on the tangent space to M_t and on the normal n, into a set of tensors on M, of orders $p, p-1, \ldots, 0$, depending on t: such decompositions have been given for the metric $\mathbf{g}$ and the Riemann tensor in Chapter 6. We say that the p-tensor u on V_T is in $E_s(T)$ if each tensor of this decomposition is in the space $E_s(T)$ of the definition. We define its norm as the sum of the corresponding norms.

2. *We define $\overset{0}{E}_s(T)$ to be the completion in the norm $E_s(T)$ of space of tensor fields in $C_0^s(T)$.*

The spaces $E_s(T)$ and $\overset{0}{E}_s(T)$ enjoy the same embedding and multiplication properties as the spaces H_s and $\overset{0}{H}_s$.

We denote by $u^{(p)}$ the set of covariant derivatives of order p of the tensor u in the metric $\mathbf{e}$.

A straightforward calculation shows that if u satisfies the system (3.2) the tensor field $u^{(p)}$ satisfies a second-order quasidiagonal system of the form

$$g^{\alpha\beta}D^2_{\alpha\beta}u^{(p)} + h_p - mu^{(p)} = f^{(p)} \tag{3.23}$$

where h_p is a linear expression in the derivatives of u of order $\leq p+1$ with coefficients derivatives of order less or equal to p of $\mathbf{g}$, b, and a together with Riemann($\mathbf{e}$) and its derivatives of order less than or equal to $p-1$, of the type (if u is a tensor field, if it is a scalar there is no $u^{(0)} \equiv u$ in the last term)

$$h_p \equiv \sum_{k=1}^{p} \mathbf{D}^{p+1-k}\mathbf{g}\mathbf{D}u^{(k)} + \sum_{k=0}^{p}\{\mathbf{D}^{p-k}a\mathbf{D}u^{(k)} + \mathbf{D}^{p-k}bu^{(k)}\}$$
$$+ \sum_{k=0}^{p} \mathbf{D}^{p-k}Riemann(\mathbf{e})u^{(k)}.$$

Lemma 3.6 *If $\mathbf{g} \in \bar{C}^0(T)$, and if $\mathbf{Dg} \in E_{s-1}(T)$ as well as a and b with $s > \frac{n}{2}+1$, then $h_p \in E_0(T)$ when $u \in E_{p+1}(T)$, for each $1 \leq p \leq s-1$ and admits the following bound for each $t \in [0,T]$*

$$||h_p||_{H_0} \leq C(T)||u||_{H_{p+1}} \tag{3.24}$$

Proof The expression of h_p and the multiplication properties of the spaces H_s. □

Theorem 3.7 *($p+1$ energy inequality) Under the hypothesis of the above lemma, and if $\boldsymbol{g}$ satisfies the Fundamental Regularity Hypothesis 2.24(1)[22] a solution $u \in \overset{0}{E}_{p+1}(T)$ of the system $Lu = f$ satisfies an inequality of the following type*

$$\| u \|_{E_{p+1}(T)} \leq C(T)\{y^{(p)}(0) + \| f \|_{E_p(T)}\}, \quad p+1 \leq s \tag{3.25}$$

where $C(T)$ is a continuous increasing function of T, depending on the norms of the coefficients of L in the functional spaces to which they have been supposed to belong and we have set

$$y^{(p)}(t) := \int_{M_t} \sum_{0\leq k\leq p+1} | \mathbf{D}^k u |^2 \, \mu_e. \tag{3.26}$$

[22] The condition $\mathbf{Dg} \in E_{s-1}(T)$, $s > \frac{n}{2}+1$ implies that $\mathbf{Dg} \in \bar{C}^0(T)$, hence 2.24(2).

Proof To make the proof easier to follow we first consider the case $p = 1$. The tensor field $u^{(1)} \equiv (D_\lambda u)$ satisfies a system of the form

$$g^{\alpha\beta} D^2_{\alpha\beta} u^{(1)} - m u^{(1)} = f^{(1)} + h_1 \tag{3.27}$$

namely:

$$g^{\alpha\beta} D^2_{\alpha\beta} D_\lambda u - m D_\lambda u = D_\lambda f + h_\lambda \tag{3.28}$$

with

$$h_\lambda \equiv a^\alpha D_\alpha D_\lambda u + D_\lambda g^{\alpha\beta} D_\alpha D_\beta u + D_\lambda a^\alpha D_\alpha u + b D_\lambda u + D_\lambda b u, \tag{3.29}$$

hence, if $s > \frac{n}{2}$, with C a Sobolev constant of (M, e) and $C(T)$ a number depending only on the bounds of the coefficients of L, it holds on each M_t, $t \in [0, T]$ that:

$$||h_1||_{H_0} \leq C(||\mathbf{Dg}||_{H_{s-1}} + ||a||_{H_{s-1}} + ||b||_{H_{s-2}})||u||_{H_2} \leq C(T)||u||_{H_2}; \tag{3.30}$$

The system for $u^{(1)}$ is of the type (3.2) with the space $\mathcal{V}$ replaced by $T^* \otimes \mathcal{V}$. If $a, \mathbf{Dg}, b \in E_{s-1}(T)$ and $u \in E_2(T)$ its coefficients satisfy the hypothesis of the first integral energy inequality. This inequality applied to the tensor $u^{(1)} \in \overset{0}{E}_1(T)$ gives:

$$z^{(1)}(t) \leq C_0(T) \left\{ z^{(1)}(0) + \int_0^t [C(T) y^{(1)}(\tau) + \parallel f^{(1)}(., t) \parallel_{H_0} z^{(1)}(\tau)^{\frac{1}{2}}] d\tau \right. \tag{3.31}$$

with

$$z^{(1)}(t) \equiv \int_{M_t} \{ \mid \mathbf{D} u^{(1)} \mid^2 + \mid u^{(1)} \mid^2 \} \mu_e$$

hence

$$y^{(1)}(t) = y(t) + z^{(1)}(t). \tag{3.32}$$

Using this inequality in conjunction with the first estimate gives

$$\begin{aligned} y^{(1)}(t) \leq C_0(T) \left\{ y^{(1)}(0) + \int_0^t [C(T) y^{(1)}(\tau) \right. \\ \left. + (||f(., ||_{H_0} + \parallel f^{(1)}(., t) \parallel_{H_0}) y^{(1)}(\tau)^{\frac{1}{2}}] d\tau \right\} \end{aligned} \tag{3.33}$$

We deduce from this integral inequality, as in previous proofs, an algebraic inequality:

$$y^{(1)}(t)^{\frac{1}{2}} \leq C_0(T) \left\{ y^1(0)^{\frac{1}{2}} + C(T) \int_0^t (\parallel f(., \tau) \parallel_{H_0} + \parallel f^{(1)}(., \tau) \parallel_{H_0}) d\tau. \right\} \tag{3.34}$$

We have thus proved the announced inequality for $p = 1$. An induction argument proves it for all $p \leq s - 1$. □

Remark 3.8 *When $m = b = 0$ we can obtain, as in the first estimate, an inequality which does not involve u itself, but only its derivatives. Under more restrictive conditions on a and b (cf. Cagnac and Choquet-Bruhat, 1984), one can then replace the condition $\varphi \in \overset{0}{H}_s$ by various hypotheses which are less restrictive in the case of non-compact manifolds. This extension is important for physical fields which do not decay fast enough in space to be square integrable, in particular for the metric itself.*

Remark 3.9 *When $u(.,0) = \varphi \in \overset{0}{H}_{p+1}$ and $\partial_t u(.,0) = \psi \in \overset{0}{H}_p$ are given the equation $Lu = f$ determines the trace on M_0 of all the derivatives of u of order $\leq p+1$, hence the function $y^{(p)}(0)$. The trace of the derivatives of order q will be in H_{p+1-q}. They will be in $\overset{0}{H}_{p+1-q}$, even when they are of order greater than 1 in t, if $f \in \overset{0}{E}_{p+1-q}$.*

The restriction pointed out in the above remark can be avoided by using the following corollary of the previous theorem.

Corollary 3.10 *Suppose that* $\mathbf{g}$ *satisfies the fundamental regularity hypothesis and* $\mathbf{Dg} \in C^0([0,T], H_{s-1})$ *as well as a and b with $s > \frac{n}{2} + 1$. Set*

$$Y^{(p)}(t) := ||u(t,.)||^2_{H_{p+1}} + ||\partial_t u(t,.)||^2_{H_p}. \tag{3.35}$$

then the following inequality holds for solutions of (3.2), $u \in C^0([0,T], H_{p+1}) \cap C^1([0,T], H_p)$, $p + 1 \leq s$.

$$Y^{(p)}(t)^{\frac{1}{2}} \leq C_0(T) \left\{ Y^{(p)}(0)^{\frac{1}{2}} + C(T) \int_0^t \| f(.,\tau) \|_{H_p} \, d\tau \right\}. \tag{3.36}$$

Proof We differentiate in space directions the system equivalent to (3.16), but with coefficient -1 for D^2_{00}

$$N^{-2}(g^{\alpha\beta} D^2_{\alpha\beta} u + a^\alpha D_\alpha u + (b - m)u) = F \equiv N^{-2} f. \tag{3.37}$$

We denote by a capital letter the product by N^{-2} of previous coefficients. We have $G^{00} = -1$. The first derivative is

$$G^{\alpha\beta} D^2_{\alpha\beta} D_i u - M D_i u = D_i F + H_i, \quad M \equiv N^{-2} m, \tag{3.38}$$

and

$$H_i \equiv D_i G^{\alpha j} D^2_{\alpha j} u + A^\alpha D_\alpha D_i u + D_i A^\alpha D_\alpha u + B D_i u + D_i (B - M) u + R_{(i)} \tag{3.39}$$

where $R_{(i)}$ is of the form

$$R_{(i)} \sim DRiemann(e)u + Riemann(e)Du. \tag{3.40}$$

It is easy to write (3.38) as a quasidiagonal second-order system for the space derivatives of u and, by methods given in the proof of the theorem, obtain the corollary. □

3.6.2 *Existence*

The existence of solutions in our spaces $\overset{0}{E}_s$ can eventually be deduced from the Cauchy–Kowalevski theorem, approximation by analytic functions and energy estimates to prove the convergence. Such a method has been used by Leray in the case of the manifold R^{n+1}, also for higher order hyperbolic equations, through the use of a global Cauchy–Kovalewski theorem. The same method applies for the spaces $C^1([0,T],\overset{0}{H}_s) \cap C^0([0,T],\overset{0}{H}_{s-1})$.

The existence of solutions in these spaces can also be deduced directly from the energy inequalities by functional methods initiated by Garding. We sketch here this method.

One proves first that the energy inequalities extend to negative values of p. One denotes[23] by H_{-p} the dual of $\overset{0}{H}_p$ (recall that $\mathcal{D}(M)$ is dense in $\overset{0}{H}_p$).

Lemma 3.11 *Suppose that $v \in \mathcal{D}(V_T)$, with $V_T \equiv [0,T] \times M$ and L is such that the p-energy estimate holds. Then the following inequality holds*

$$Sup_{0\leq t\leq T} \parallel v(t,.) \parallel_{H_{-p}} \leq C(T) \int_0^T \parallel Lv(t,.) \parallel_{H_{-p-1}} dt,$$

where $C(T)$ is a continuous increasing function of T depending on the coefficients of L.

Proof See for instance Sogge[24]. □

Theorem 3.12 *(Existence in a strip)*

1. *If the metric $\boldsymbol{g}$ and the coefficients a, b satisfy the hypothesis of Corollary 3.10 and if we suppose that $f \in L^1([0,T],\overset{0}{H}_p)$, and the Cauchy data are $(\varphi,\psi) \in \overset{0}{H}_{p+1} \times \overset{0}{H}_p$, then the equation $Lu = f$ has one and only one solution on $V = M \times R$ taking these Cauchy data such that for any $T \in R$*

$$u \in C^0([0,T],\overset{0}{H}_{p+1}) \cap C^1([0,T],\overset{0}{H}_p), \quad 0 \leq p \leq s-1. \tag{3.41}$$

The norm of u depends continuously on the norms of ϕ and ψ.

[23] CB-DM1 p. 491.
[24] Sogge, C. (1995) *Lectures on Non-Linear Wave Equations*, International Press.

2. *If the coefficients, and f, satisfy the hypothesis of Lemma 3.6, then the solution is in $\overset{0}{E}_{p+1}$.*

Proof

1. We give a proof for zero Cauchy data. The function $f \in L^1([0,T)], \overset{0}{H}_p)$ defines a linear form on $\mathcal{D}(V_T)$ by the formula:

$$\langle f, v\rangle := \int_0^T (f(t,.), v(t,.))_{H_0} dt \tag{3.42}$$

We have, by the definition of H_{-p},

$$(f(t,.), v(t,.))_{H_0} \leq \parallel f(t,.) \parallel_{H_p} \parallel v(t,.) \parallel_{H_{-p}} \tag{3.43}$$

We deduce from Lemma 3.11 ($-p$-energy inequality), with $C(T)$ a number depending on the bounds of the coefficients of L,

$$\mid \langle f, v\rangle \mid \leq C(T) \parallel f \parallel_{L^1([0,T],H_p)} \int_0^T \parallel Lv(t,.) \parallel_{H_{-p-1}} dt \tag{3.44}$$

Hence $v \mapsto \langle f, v\rangle$ defines a linear form which is continuous in the norm $L^1([0,T], H_{-p-1})$ on the subspace Lv, $v \in \mathcal{D}(V_T)$. By the Hahn–Banach theorem there exists u (non unique) in the dual, $L^\infty([0,T], \overset{0}{H}_{p+1})$, such that

$$\langle f, v\rangle = \langle u, Lv\rangle, \quad v \in \mathcal{D}(V_T)$$

hence, with L^* the L^2 adjoint of the operator L:

$$L^* u = f \tag{3.45}$$

in the sense of distributions. One completes the proof by more elaborate reasoning, using as in the first existence theorem the fact that $\langle u, Lv\rangle = \langle L^* u, v\rangle$ if u and v take Cauchy data zero respectively on $t = 0$ and $t = T$.

2. One deduces from the equation satisfied by $u \in C^0([0,T], \overset{0}{H}_{p+1}) \cap C^1([0,T], \overset{0}{H}_p)$ that $u \in C^2([0,T], \overset{0}{H}_{p-1})$ if $f \in C^0([0,T], \overset{0}{H}_{p-1})$. Derivation of the equation with respect to t, together with an induction argument, completes the proof.

□

3.7 *Other hypotheses on* **g**

Many variants of the existence theorem 3.12, adapted to the problem to solve, can be obtained. The proof of Theorem 3.12 relies on the energy estimate of Theorem 3.7, with the hypotheses of Lemma 3.6. These hypotheses can be replaced by other conditions on the metric **g** which ensure the conclusion of the lemma.

We give an example[25]. We denote by $E^p_\sigma(T)$ the following Banach space generalization of the spaces $E_\sigma(T)$:

$$E^p_\sigma(T) \equiv \cap_{0\le k\le\sigma} C^k([0,T], W^p_{\sigma-k}).$$

where W^p_s is a Sobolev, Banach, space[26] with norm

$$\| f \|_{W^p_s} := \left\{ \int_M \sum_{0\le k\le s} | D^k f |^p \, \mu_e \right\}^{1/p}$$

Lemma 3.13 *If $n = 2$ or 3 the hypotheses $\boldsymbol{g} \in \bar{C}^0$ and **Dg,** $a, b \in E^p_3(T)$, with $p > 1$, imply that $h_\sigma \in E_0(T)$ when $u \in E_\sigma(T)$ for each $1 \le \sigma \le 2$.*

Proof The hypothesis **Dg,** $a \in E^p_3(T)$, $p > 1$, implies that if $n \le 3$, then **Dg,** $a \in C^0_b(T)$. The multiplication properties of Sobolev spaces complete the proof. □

This lemma leads to the following existence theorem.

Theorem 3.14 *Let M be a compact manifold and $n \le 3$. If the coefficients of L satisfy the Fundamental Regularity Hypothesis 2.24 and the assumptions of the previous lemma, and if we suppose that $f \in L^1([0,T], H_1)$, and the Cauchy data are $(\varphi, \psi) \in H_2 \times H_1$, then the equation $Lu = f$ has one and only one solution taking these Cauchy data, $u \in C^0([0,T], H_2) \cap C^1([0,T], H_1)$.*

3.8 *Cauchy data in local spaces*

A space H_s coincides with the space $\overset{0}{H}_s$, closure in the norm H_s of the space $\mathcal{D}(M)$ of C^∞ functions with compact support in M only under some conditions recalled in Appendix I. These conditions are satisfied for all s if e is a smooth metric on a compact manifold (without boundary, as we always mean if not otherwise specified), or if (M, e) is a smooth manifold Euclidean at infinity. The existence theorem that we have proved on a strip restricts, if M is not a compact manifold (nor (M, e) Euclidean at infinity), the behaviour in space not only of the Cauchy data and of the solution as in Theorem 2.29, but also, if $s \ge 2$, of the coefficient f of the equations.

In the case where (M, e) can be completed by a boundary ∂M we have solved in fact a mixed initial and boundary values problem: Cauchy data on M and zero boundary values on ∂M.

In the case where (M, e) is complete with curvature satisfying the relevant boundedness hypothesis we have already indicated that the hypothesis

[25] Used in the proof of global existence of solutions of Einstein equations with U(1) isometry group by Choquet-Bruhat and Moncrief (2001); see more comments and references in Chapter 16.

[26] See Appendix A.

$\varphi \in \overset{0}{H}_s \equiv H_s$ on the Cauchy data φ, ψ can be replaced, in the case $m - b = 0$, by $\varphi - \varphi_0 \in H_0$, with φ_0 some given function, while $D\varphi, \psi \in H_{s-1}$, properties possibly compatible with physical behaviour on an asymptotically Euclidean three-dimensional manifold.

Though we have proved our existence theorem in a strip, we know by the local uniqueness and causality results that the local in time behaviour of our hyperbolic system does not depend on the global in space behaviour of the Cauchy data. We will prove theorems which do not imply any decay of the data.

3.8.1 *Uniform spaces*

Definition 3.15 *A tensor field u on (M, e) is in $H_{s,loc}$ if it is in $H_s(U, e)$ for any relatively compact open subset U of M.*

The space $H_{s,loc}$ is neither a Hilbert nor a Banach space. One can define a Banach space of tensor fields on (M, e) as follows.

Definition 3.16 *Let $U_{(i)}$ be a locally finite covering of M by relatively compact open sets $U_{(i)}$. We say that a tensor field $u \in H_{s,loc}$ is in $H_{s,u.l}$ if:*

$$\| u \|_{H_{s,u.l}} \equiv Sup_i \| u \|_{H_s(U_{(i)})} < \infty.$$

The space $H_{s,u.l}$ is a Banach space.

Remark 3.17 *$H_{s,u.l}$ is a subspace of $H_{s,loc}$, because any compact set of M can be covered by a finite number of the $U'_{(i)}s$. On a compact manifold a space $H_{s,loc}$ is also a space $H_{s,u.loc}$, but if M is not compact, this is no more the case.*

The space $H_{s,u.l}$ has the embedding and multiplication properties recalled in Appendix A if the manifolds $(U_{(i)}, e)$, are Sobolev regular. We will say that (M, e) is u.l. (uniformly locally) Sobolev regular if it admits such a covering and the Sobolev constants relative to these various subsets are uniformly bounded[27]. It will be the case in particular if (M, e) has a non-zero injectivity radius d: we can then choose for $U_{(i)}$ geodesic balls of radius $d > 0$, they are Sobolev regular[28]; we can also consider a locally finite atlas of M where the images of domains of charts are balls of R^n with fixed radius, and the restriction of the representative of e in these balls is uniformly equivalent to the Euclidean metric.

We define the spaces $E_{s,u.l}$ analogously to $E_s(T)$, replacing H_k by $H_{k,u.l}$.

We can now prove our existence theorem in local spaces. For simplicity of writing, and in view of application to non-linear equations we take[29] $a = b = m = 0$.

[27] It is not the case if their e-radius is not uniformly bounded below by a strictly positive number.

[28] Recall that (M, e) is supposed to be smooth.

[29] These coefficients are then included in the right-hand side.

Theorem 3.18 *Let*

$$Lu \equiv g^{\alpha\beta} D^2_{\alpha\beta} u = f$$

be a linear quasidiagonal hyperbolic system on the regularly sliced (hypothesis 2.24) hyperbolic manifold $(V = M \times R, \boldsymbol{g})$. *Suppose that* (M, e) *is u.l. Sobolev regular for the covering* $(U_{(i)}, e)$, *and that* $\boldsymbol{Dg}$, $f \in C^0([0,T], H_{s-1,u.l}]$, $s > \frac{n}{2} + 1$. *Then the Cauchy problem with data* $(\varphi, \psi) \in H_{\sigma,u.l} \times H_{\sigma-1,u.l}$ *on* M, $1 \leq \sigma \leq s$ *has a global solution* u *on* V. *It has a restriction to each* M_t *such that*

$$u(.,t) \in H_{\sigma,u.l}, \quad \partial_t u(.,t) \in H_{\sigma-1,u.l}$$

Proof Consider a locally finite covering $\{U_{(i)}, e\}$ of a slice M_T. Take T small enough for the existence of basic (see Section 2.4) compact causal domains $\Omega_{(i)}$, with bases $S_{(i),0}$ covering M_0 and upper boundaries $S_{(i),T}$ covering M_T. The Cauchy problem for our linear system with data on M_0 in $\overset{0}{H}_{\sigma} \times \overset{0}{H}_{\sigma-1}$ coinciding with (φ, ψ) on $S_{(i)}$ has a solution $u(i)$ on V, which depends only in $\Omega_{(i)}$ on the restriction of (φ, ψ) to $S_{(i)}$. Moreover solutions so defined in two compact causal domains $\Omega_{(i)}$ and $\Omega_{(j)}$ coincide in $\Omega_{(i)} \cap \Omega_{(j)}$ by the local uniqueness and causality theorem, since they take the same Cauchy data on $S_{(i)} \cap S_{(j)}$. The set of restrictions of the $u_{(i)}$ to the $\Omega_{(i)}$ defines therefore a field u on V_T, a solution of the given Cauchy problem. Repeating this operation for successive intervals of t completes the proof of the theorem. □

Other variants of the functional spaces for which we have proved existence theorems can evidently be considered.

Example. On a Lorentzian manifold $(V, \mathbf{g})$ with V the subspace of R^4 given by $x^1 - t > 0$ one may want to consider Cauchy data which are in the closure in H_s norms of functions with compact support in $[0, \infty) \times R^2$. Energy type estimates may be obtained using the fact that the boundary of V as a subset of R^4 is a null submanifold. The problem considered in this example arises in particular when the Lorentzian manifold is the exterior of a black hole.

3.8.2 *Local solutions*

The uniqueness and causality theorem (in Section 3.5) leads to the existence of local in space solutions, with Cauchy data on an open set U of a spacelike submanifold S. Such solutions can be obtained as restrictions to a compact causal domain with basis $U \subset S$ of a global in space solution with Cauchy data obtained by appropriate prolongation to S of the data given on U. Such a procedure permits, due again to the uniqueness and causality theorem, constructions of solutions of the Cauchy problem for the PDE considered in this section in a neighbourhood of an initial submanifold which is not part of a regular slicing. The solution is not necessarily global on the Lorentzian manifold $(V, \mathbf{g})$, even for linear systems[30].

[30] See more comments in Chapter 12.

4 Quasilinear systems

We consider general quasilinear quasidiagonal second-order systems,

$$\mathbf{g}(.,u,\mathbf{D}u).\mathbf{D}^2u + f(.,u,\mathbf{D}u) = 0. \tag{4.1}$$

where u is some tensor (or spinor) field over V and $\mathbf{D}u$ its covariant derivative in the properly given Riemannian metric $\mathbf{e}$. The symmetric 2-tensor field $\mathbf{g}$ and the tensor field f of the same type as u are given in a chart of local coordinates through usual functions from R^P into R^N, $P = n + 1 + N + (n + 1)N$.

4.1 *Semilinear systems*

We first consider the semilinear wave equation for a scalar function u on a manifold V with $\mathbf{g}$ a given Lorentzian metric:

$$\Box_{\mathbf{g}}u \equiv \mathbf{g}.\nabla^2 u = f(.,u,\mathbf{D}u) \tag{4.2}$$

The theorem that we will prove for the semilinear wave equation extends in a straightforward way to quasidiagonal semilinear systems.

We make on (V,$\mathbf{g}$) the hypothesis made previously for general linear equations. Namely V is a product $M \times R$ endowed with a smooth proper Riemannian metric $\mathbf{e} = dt^2 + e$. The Sobolev spaces on M are defined though e, supposed Sobolev regular in the sense of section 3.6.

Hypothesis L:

1. The space time metric $\mathbf{g}$ satisfies on $V_T = M \times [0,T]$ the Fundamental Regularity hypothesis 2.24. (1)
2. $\mathbf{Dg} \in C^0([0,T],H_{s-1})$, $s > \frac{n}{2} + 1$, this hypothesis implies 2.24. (2)

On a generic function f we make the following hypothesis:

Hypothesis SL:
There exist numbers k and k' defining intervals B and B' of R, which we choose centred at zero for simplicity of writing, such that if the $\mathbf{e}$-norms of the function v and the covariant vector field w lie in these intervals, i.e. satisfy the inequalities

$$|v| < k, \quad |w| < k', \tag{4.3}$$

then the mapping $(x,v,w) \mapsto f(x,v,w)$ is a C^∞ mapping in v and w at each point x of V_T, and a $C^0([0,T],\overset{0}{H}_{s-1})$ mapping in x for each v and w.

We denote by $B_s(T)$ the Banach space, subset of $E_s(T)$, defined by

$$B_s(T) := C^0([0,T],H_s) \cap C^1([0,T],H_{s-1}) \tag{4.4}$$

Lemma 4.1 *If $s > \frac{n}{2} + 1$, there exists a ball $B_K \subset B_s(T) \equiv C^0([0,T],H_s) \cap C^1([0,T],H_{s-1})$ with radius K independent of T,*

$$\| v \|_{B_s(T)} \equiv Sup_{0\le t\le T}\{||v(t,.)||_{H_s} + ||\partial_t v(t,.)||_{H_{s-1}}\} < K, \tag{4.5}$$

such that $v \in B_K$ implies that $f(.,v,\mathbf{D}v) \in C^0([0,T],H_{s-1})$.

Proof The Sobolev embedding theorem proves that there exist constants depending only on (M, e) such that, if $s > \frac{n}{2} + 1$,

$$|v(t,.)|, |Dv(t,.)| \leq C||v(t,.)||_{H_s}, \quad \text{and} \quad |\partial_t v(t,.)| \leq C||\partial_t v(t,.)||_{H_{s-1}}. \tag{4.6}$$

The inequality (4.5) therefore implies (4.3) if $K > 0$ is appropriately chosen. The Leray–Sobolev[31] composition theorem completes the proof. □

In the case where the manifold (M, e) is not complete we make the further hypothesis:

Hypothesis: $\overset{0}{SL}$. In the case where $H_s \neq \overset{0}{H}_s$ we suppose that $v \in \overset{0}{B}_s(T) := C^0([0,T], \overset{0}{H}_s) \cap C^1([0,T], \overset{0}{H}_{s-1})$ also implies that $f(., v, \mathbf{D}v) \in C^0([0,T], \overset{0}{H}_{s-1})$.

Theorem 4.2 *(Local in time existence and uniqueness). Under the hypothesis L, SL, and* $\overset{0}{SL}$*, there exists* $\tau > 0$ *such that the equation*

$$\mathbf{g}.\nabla^2 u = f(., u, \mathbf{D}u) = 0$$

has one and only one solution $u \in \overset{0}{B}_s(\tau)$ *taking on* M_0 *the Cauchy data* $u(.,0) = \varphi, \partial_t u(.,0) = \psi$*, if*

$$\varphi \in \overset{0}{H}_s, \quad \psi \in \overset{0}{H}_{s-1}, \quad \| \varphi \|_{H_s} + \| \psi \|_{H_{s-1}} < K, \quad s > \frac{n}{2} + 1.$$

The norm of u depends continuously on the norms of ϕ and ψ.

Proof Suppose that $v \in \overset{0}{B}_s(t)$, $t \in [0,T]$. Then, by the lemma and the third hypothesis, $f(., v, \mathbf{D}v) \in C^0([0,t], \overset{0}{H}_{s-1})$ with:

$$\| f(., v, \mathbf{D}v) \|_{C^0([0,t], H_{s-1})} \leq F(T, K)$$

where $F(T, K)$ is some number depending only on K and the bound of the $C^0([0,t], H_{s-1})$ norm of the function $x \mapsto f(x, v, w)$ for fixed v and w.

The section on general linear systems tells us that the Cauchy problem with the given data (φ, ψ) for the linear equation

$$\mathbf{g}.\nabla^2 u = f(., v, \mathbf{D}v) = 0$$

has one and only one solution $u \in \overset{0}{B}_s(t)$ for each $t \in [0,T]$. It satisfies the following inequality:

$$\| u \|_{B_s(t)} \leq C_0(t)\{\| \varphi \|_{H_s} + \| \psi \|_{H_{s-1}}\} + tC(T)F(T, K)$$

[31] See Dionne, P. (1962) *J. d'Anal. Math. Jérusalem*, **10**, 1–90. The spaces B_s and $(C^0[0,T], H_{s-1})$ can be replaced by E_s and E_{s-1}.

with

$$C_0(t) \leq C_0(0) + tC(T). \tag{4.7}$$

We denote by $\mathcal{F}_t$ the mapping of $\overset{0}{B}_s(t)$ into the Banach space $\overset{0}{B}_s(t)$ defined by associating with v this solution u of the linear problem. We shall show that there exists $\tau > 0$ such that the mapping $\mathcal{F}_\tau$ is contracting from a closed ball into itself, hence admits a fixed point.

By the hypothesis on the Cauchy data there exists K_0 such that

$$\| \varphi \|_{H_s} + \| \psi \|_{H_{s-1}} < K_0 < K. \tag{4.8}$$

There exists a number τ such that

$$\| u \|_{B_s(\tau)} < K, \tag{4.9}$$

i.e. u lies in the ball B_K, if

$$\tau < \frac{K - C_0(0)K_0}{C(T)(\| \varphi \|_{H_s} + \| \psi \|_{H_{s-1}} + F(T,K)}. \tag{4.10}$$

We have $\tau > 0$ if

$$K > C_0(0)K_0. \tag{4.11}$$

This inequality reduces to $K > K_0$ if $C_0(0) = 1$, property which we achieve following Lemma 2.25 by choosing $\mathbf{e} = \mathbf{G}_{\nu,\nu}|_{\mathbf{t=0}}$.

To show that the mapping is contracting we consider two elements v_1 and v_2 of that ball and the corresponding images u_1 and u_2. They are such that

$$\mathbf{g}.\nabla^2(u_1 - u_2) = f(., v_1, \mathbf{D}v_1) - f(., v_2, \mathbf{D}v_2), \tag{4.12}$$

and $u_1 - u_2$ has zero Cauchy data. Using again the energy inequality, and the hypothesis on f we obtain an inequality of the form

$$\| u_1 - u_2 \|_{B_s(t)} \leq C(T)F(T,K)t \| v_1 - v_2 \|_{B_s(t)} . \tag{4.13}$$

We see on this inequality that we can again choose $0 < \tau \leq T$ small enough for the mapping $\mathcal{F}_\tau$ to be contracting, namely

$$\tau < \frac{K}{C(T)F(T,K)}.$$

The inequalities restricting the time τ of existence of the solution show that this time depends only on the fundamental bounds of the coefficients and of the norms of the Cauchy data. Inequality (4.13) also proves the uniqueness of the solution; and continuous dependence on data. □

4.2 *Further results for semilinear equations*

4.2.1 *Other functional spaces*

Modulo further hypotheses on the time derivatives of $\mathbf{g}$ and f, if f depends explicitly directly on t, one obtains, along the same lines, a local in time existence theorem in the space $\overset{0}{E}_s$.

One can also obtain existence theorems on a strip in the uniformly local Sobolev spaces, and existence theorems in compact causal domains, as we did for linear systems.

4.2.2 *Special equations*

We have considered the case of a generic non-linear source f. In many cases of physical interest the source f enjoys special properties which make the local in time existence result hold in spaces involving a smaller s: the proof by iteration of Theorem 4.2 shows that the existence result holds in $\overset{0}{E}_s$ if $f(u, \mathbf{D}u) \in \overset{0}{E}_{s-1}$ when $u \in \overset{0}{E}_s$.

Examples.

(a) The power law wave equation, with P a positive integer, $m \geq 0$:

$$\mathbf{g} \cdot \nabla^2 u - mu = u^P \tag{4.14}$$

The source f does not depend on $\mathbf{D}u$ and is C^∞ in u for all u.

The Sobolev embedding theorem shows that $u^P \in E_0(T)$ as soon as $u \in \overset{0}{E}_1$ if

$$P \leq \frac{n}{n-2} \tag{4.15}$$

We obtain a local[32] existence and uniqueness theorem for Cauchy data in $\overset{0}{H}_1 \times H_0$, if $\mathbf{g}$ satisfies the Fundamental Regularity Hypothesis 2.24.

Inequality (4.15) also implies $u^P \in \overset{0}{E}_{s-1}$ as soon as $u \in \overset{0}{E}_s$ for any p. The Cauchy problem for (4.14) has one and only one solution for Cauchy data in $\overset{0}{H}_s \times \overset{0}{H}_{s-1}$, modulo regularity hypothesis on $\mathbf{g}$ which always include the Fundamental Regularity Hypothesis.

(b) Equation

$$\mathbf{g} \cdot \nabla^2 u = u \partial_0 u \tag{4.16}$$

The multiplication theorem $H_s \times H_{s-1} \to H_{s-1}$ when $s > \frac{n}{2}$ shows that the local existence and uniqueness result holds for Cauchy data in $\overset{0}{H}_s \times \overset{0}{H}_{s-1}$, $s > \frac{n}{2}$.

[32] Some global existence theorems can also be proved, in particular if $n = 3$, $P = 3$.

The Yang–Mills equations on a four-dimensional spacetime involve terms of the kind given in the above examples, they admit a local in time solution for Cauchy data in $\overset{0}{H}_2 \times \overset{0}{H}_1$.[33]

Wave maps on two- or three-dimensional spacetimes can also be proved, using interpolation theorems, to have local in time solutions for $\overset{0}{H}_2 \times \overset{0}{H}_1$ Cauchy data.

4.3 *Quasilinear systems*

When the metric $\mathbf{g}$ depends on u, and possibly on $\mathbf{D}u$, with u a function or a set of scalar functions or a tensor or spinor field over V, the system

$$\mathbf{g}(., u, \mathbf{D}u).\mathbf{D}^2 u + f(., u, \mathbf{D}u) = 0.$$

is called quasilinear.

4.3.1 *Existence and uniqueness theorem*

We will make on f the hypotheses SL of Section 4.1 (semilinear case), and on $\mathbf{g}$ hypotheses which ensure that the symmetric 2-tensor field $x \mapsto \mathbf{g}(x, u(x), \mathbf{D}u(x))$ satisfies the hypothesis L, at least in the neighbourhood of some field u_0 on V, which we choose to be zero for simplicity of writing. In order to avoid restrictions "at infinity" on non-compact manifolds for the field $\mathbf{g}$ and to write more concisely we introduce the following definition for tensor fields on $V_T \equiv (M \times [0,T], \mathbf{e}]$.

Definition 4.3 *We denote by $\tilde{B}_s(T)$ the Banach space of tensor fields v on V_T defined by*

$$\tilde{B}_s(T) := \{v \in \bar{C}^0(T), \quad \mathbf{D}v \in C^0([0,T], \overset{0}{H}_{s-1})$$

We replace $\overset{0}{B}_s$ by $\tilde{B}_s$ in the hypothesis[34] SL.

Hypothesis NL

There are numbers k and k' such that if

$$\| v \|_{\bar{C}^0(T)} < k, \quad \| w \|_{\bar{C}^0(T)} < k' \tag{4.17}$$

then

1(a). The symmetric 2-tensor field $x \mapsto \mathbf{g}(x, v(x), w(x))$ satisfies the Fundamental Regularity Hypothesis 2.24.

1(b). $\mathbf{g}(x, v, w)$ is C^∞ in v and w for each $x \in V_T$, and is in $\tilde{B}_{s-1}(T)$ in x for each fixed pair v, w.

2. the mapping $(x, v, w) \mapsto f(x, v, w)$ is a C^∞ mapping in v and w at each point x of V_T, and a $C^0([0,T], \overset{0}{H}_{s-1})$ mapping in x for each v and w.

[33] Klainerman and Machedon have proved by more refined methods the existence in "energy space", that is for Cauchy data in $H_1 \times H_0$.

[34] Other choices can be made, adapted to the problem to solve.

Lemma 4.4 *There exists a ball B_K of the Banach space $\tilde{B}_s$,*

$$\| u \|_{\tilde{B}_s} < K,$$

such that if $u \in B_K$, then the tensor field $x \mapsto \mathbf{g}(x, u(x), \mathbf{D}u(x))$ is in $\tilde{B}_{s-1}(T)$ and satisfies the Fundamental Regularity Hypothesis 2.24(1) if $s > \frac{n}{2} + 1$. It also satisfies hypothesis 2.24(2), $\boldsymbol{Dg} \in C^0([0,T), H_{s-2}) \subset \bar{C}^0(T)$, if $s > \frac{n}{2} + 2$.

Proof Analogous to the proof of Lemma 4.1. □

Remark 4.5 *The system (4.1) is called **hyperquasilinear** if the metric $\boldsymbol{g}$ depends on u but not $\boldsymbol{D}u$. In the hyperquasilinear case, under the hypothesis spelled out above, the metric tensor $\boldsymbol{g}$ belongs to $\tilde{B}_s$ when $u \in \tilde{B}_s$, as soon as $s > \frac{n}{2} + 1$ (and then $\boldsymbol{Dg} \in \bar{C}^0$).*

The Einstein equations in harmonic gauge are a hyperquasilinear system whose coefficients satisfy the hypotheses NL and SL.

We are now ready to prove the following theorem, where $\tilde{H}_s$ denotes the space of tensor fields φ on (M, e) such that $\varphi \in \bar{C}^0$ and $D\varphi \in \overset{0}{H}_{s-1}$. The norm is

$$\| \varphi \|_{\tilde{H}_s} \equiv \| \varphi \|_{\bar{C}^0} + \| D\varphi \|_{H_{s-1}} .$$

Theorem 4.6 *(Local in time existence). Under the hypotheses NL, SL, and $\overset{0}{SL}$, there exists $\tau > 0$ such that the hyperquasilinear system*

$$\mathbf{g}(., u).\mathbf{D}^2 u + f(., u, \mathbf{D}u) = 0. \tag{4.18}$$

has a solution $u \in \tilde{B}_s(\tau)$ taking on M_0 the Cauchy data

$$u(., 0) = \varphi \in \tilde{H}_s, \quad \partial_t u(., 0) = \psi \in \overset{0}{H}_{s-1}, \tag{4.19}$$

if $s > \frac{n}{2} + 1$ and the Cauchy data are such that

$$\| \varphi \|_{\bar{C}^0} < k, \quad \| D\varphi \|_{\bar{C}^0} + \| \psi \|_{\bar{C}^0} < k'. \tag{4.20}$$

Proof One constructs as in the semilinear case a mapping $\mathcal{F}_\tau : v \mapsto u$ from $\tilde{\mathcal{B}}_s(\tau)$ into itself, for τ small enough, by solving the linear system

$$\mathbf{g}(., v).\mathbf{D}^2 u = f(., v, \mathbf{D}v);$$

however, there appears a difficulty in proving the contracting property of the mapping $\mathcal{F}_\tau$: if u_1 and u_2 are respectively the solutions of the following systems:

$$\mathbf{g}^{\alpha\beta}(., v_i)\mathbf{D}^2_{\alpha\beta} u_i + f(., v_i, \mathbf{D}v_i) = 0, \quad i = 1 \text{ or } 2$$

the difference $u_1 - u_2$ satisfies a system given by:

$$g^{\alpha\beta}(., v_1) D^2_{\alpha\beta}(u_1 - u_2) + \{g^{\alpha\beta}(., v_1) - g^{\alpha\beta}(., v_2)\} D^2_{\alpha\beta} u_2 + f(., v_1, \mathbf{D}v_1) - f(., v_2, \mathbf{D}v_2) = 0.$$

This system, due to the presence of the term $D^2_{\alpha\beta}u_2$, has a source which is not in $\overset{0}{B}_{s-1}$ when v_i and hence u_i are in $\tilde{B}_s$, However this source is always in $\overset{0}{B}_{s-2}$ when $s-1 > \frac{n}{2}$. We can apply the previous method to show that the mapping $\mathcal{F}_\tau$: $v \mapsto u$ from $\tilde{\mathcal{B}}_s$ into itself is contracting, but now for the norm B_{s-1}. Returning to the formulation of the contracting mapping principle as the convergence of an iteration scheme we obtain the existence, for τ small enough, of a function $u \in \tilde{B}_{s-1}(\tau)$, limit of a sequence of functions $u_n \in \tilde{B}_s(\tau)$, such that $u_n = \mathcal{F}_\tau u_{n-1}$. We know moreover that the functions Du_n are uniformly bounded in $B_{s-1}(\tau)$ norm, and hence *a fortiori* in the norm $L^\infty([0,\tau], H_{s-1})$. This space is the dual of the Banach space $L^1([0,\tau], H_{s-1})$ for the duality defined by the following integral where $(.,.)_{H_{s-1}}$ denotes the scalar product in the Hilbert space H_{s-1}:

$$\langle f, h\rangle = \int_0^\tau (f(t,.), h(t,.))_{H_{s-1}} dt$$

Therefore there exists a subsequence of $\{Du_n\}$ which converges weakly, and hence in particular in the sense of distributions, towards $u' \in L^\infty([0,\tau], H_{s-1})$. Since the limit of a sequence is unique in $\mathcal{D}'$ this u' is the preceding Du. We prove that the obtained solution u is indeed in $\tilde{E}_s(T)$ through the following lemma. □

Lemma 4.7

1. *The function* $Du \equiv v : t \mapsto Du(t,.)$ *is continuous from* $[0,\tau]$ *into* $\overset{0}{H}_{s-1}$ *for the weak topology on* $\overset{0}{H}_{s-1}$ *defined by its duality with* H_{-s+1}, *i.e. the real-valued function:*

$$t \mapsto \langle v(t,.), \phi\rangle$$

is continuous for any $\phi \in H_{-s+1}$.

2. $Du \equiv v \in C^0([0,\tau], \overset{0}{H}_{s-1})$.

Proof We consider a sequence $\phi_n \in H_{-s+2}$ converging to ϕ in H_{-s+1}, while $v_n(t,.)$ converges uniformly to v in $\overset{0}{H}_{s-2}$ and is uniformly bounded in $\overset{0}{H}_{s-1}$. We prove that $< v_n, \phi >$ converges uniformly to $< v, \phi >$ for any $\phi \in H_{-s+1}$ by writing

$$\begin{aligned} |\,(v,\phi) - (v_n,\phi)\,| \leq{}& \|\, v - v_n \,\|_{H_{s-2}} \|\, \phi_n \,\|_{H_{-s+2}} \\ &+ (\|\, v \,\|_{H_{s-1}} + \|\, v_n \,\|_{H_{s-1}}) \,\|\, \phi - \phi_n \,\|_{H_{-s+1}} \end{aligned}$$

We conclude that $t \mapsto v(t,.)$ is, like $t \mapsto v_n(t,.)$, continuous in the weak topology of $\overset{0}{H}_{s-1}$.

2. It is sufficient to prove continuity on $[0,\tau]$ to prove continuity at some point of this interval. Moreover, the right continuity implies continuity due to time reversibility. The right continuity at $t = 0$, in the norm H_{s-1}, of the function

$v \mapsto v(t,.)$ will be a consequence of its weak continuity if the norm of its weak limit is such that:

$$\| v(0,.) \|_{H_{s-1}} \geq LimSup_{t \mapsto 0, t \geq 0} \| v(t,.) \|_{H_{s-1}}$$

This inequality can be proved through detailed consideration of the previous proof of the s-energy inequality. □

Corollary 4.8 *The solution is unique and depends continuously on the data.*

Proof The difference of two solutions u_1 and u_2 satisfies Equation (4.17) (with $u_1 = v_1$ and $u_2 = v_2$)

$$g^{\alpha\beta}(.,u_1)D^2_{\alpha\beta}(u_1 - u_2) + \{g^{\alpha\beta}(.,u_1) - g^{\alpha\beta}(.,u_2)\}D^2_{\alpha\beta}u_2 + f(.,u_1,\mathbf{D}u_1) - f(.,u_2,\mathbf{D}u_2) = 0. \tag{4.21}$$

The mean value theorem implies, under the functional hypothesis, pointwise estimates of the form

$$|g^{\alpha\beta}(.,u_1) - g^{\alpha\beta}(.,u_2)| \leq C|u_1 - u_2| \tag{4.22}$$

$$|f(.,u_1,\mathbf{D}u_1) - f(.,u_2,\mathbf{D}u_2)| \leq C\{|u_1 - u_2| + |\mathbf{D}u_1 - \mathbf{D}u_2|\} \tag{4.23}$$

The first energy inequality proved for linear systems is sufficient to prove the uniqueness of the solution for given Cauchy data.

The proof of the continuous dependence of the solution in $\tilde{B}_s$ on the Cauchy data in $\tilde{H}_s \times H_{s-1}$ requires a recursive argument. □

Corollary 4.9 *In the quasilinear case, the theorem and its corollary hold for* $s > \frac{n}{2} + 2$.

Proof Straightforward, using analogous arguments. □

Corollary 4.10 *When the direct time dependence of* $\boldsymbol{g}$ *and* f *is adequate, in particular when there is no such dependence, the existence holds in* $\tilde{E}_{s(\tau)}$.

4.3.2 *Smooth solutions*

We denote by s_0 the smallest integer greater than $\frac{n}{2} + 2$ (respectively $\frac{n}{2} + 1$ in the hyperquasilinear case). We have shown that there exists $\tau > 0$ such that the Cauchy problem for the quasilinear system with data (φ, ψ) in appropriate balls of the space $\tilde{H}_{s_0} \times H_{s_0-1}$ has a solution on V_τ, $u \in \tilde{B}_{s_0}(\tau)$.

The following theorem is important, particularly for its corollary.

Theorem 4.11 *If in addition to the previous hypothesis the Cauchy data belong to* $\tilde{H}_s \times H_{s-1}$ *with* $s > s_0$, *then the solution* u *is in fact in* $\tilde{B}_s(\tau)$ *possibly in* $\tilde{E}_s(\tau)$, *with the same* τ *as for* $s = s_0$.

Corollary 4.12 *If the initial data* (φ, ψ) *are in* C^∞ *the solution is* C^∞ *on* V_τ.

Proof Let us consider first hyperquasilinear systems and, for simplicity of writing suppose that neither $\mathbf{g}$ nor f depend directly on the point $x \in V$. It is easy to extend the results to the case where they do by making appropriate hypotheses on the E_s regularity of this dependence.

By differentiating the given system we obtain a quasidiagonal second-order system for $\mathbf{D}u \equiv U$ of the form

$$g(u).\mathbf{D}^2U + g'_uU\mathbf{D}U + g(u, \mathbf{D}u)Riemann(\mathbf{e})U = f'_uU + f'_{\mathbf{D}u}\mathbf{D}U$$

The above system is linear in U when u is given. If $u \in \tilde{B}_{s_0}(\tau)$ with Cauchy data in the given balls this system for U satisfies the hypothesis of the existence and uniqueness theorem for linear systems with $s = s_0$, it has therefore a solution U on V_τ which is in $\tilde{B}_{s_0}(\tau)$, and coincides with $\mathbf{D}u$, hence $u \in \tilde{E}_{s_0+1}(\tau)$.

A recursive argument can be applied in the case of a fully quasilinear system, after twice differentiating the original system.

The corollary results from the Sobolev embedding theorem: a solution which is in $E_s(\tau)$ for all $s > s_0$ is in $C^\infty(\tau)$. □

4.3.3 *Causality. Existence in local spaces*

We proved in Section 3.3 an energy inequality for linear systems in a compact causal domain, and we have deduced from it a causality statement. An analogous property holds in the quasilinear case.

Theorem 4.13 *(Causality) Let u_1 and u_2 be two solutions in $\tilde{B}_s(\tau)$ of a quasilinear system which satisfy the hypotheses of the existence theorem 4.6. Suppose these two solutions take the same Cauchy data on the basis of a compact subset Ω causal for the Lorentzian metric $\boldsymbol{g}_1$ determined by u_1. Then $u_2 = u_1$ in Ω.*

Proof The energy inequality in the domain Ω applied to the equation satisfied by the difference of two solutions, as in the proof of uniqueness in Section 4.1. □

It is then straightforward to prove for quasilinear systems a theorem analogous (but now local in time) to the existence theorem 3.18 proved for linear systems

Theorem 4.14 *(Existence in local spaces) A quasilinear quasidiagonal system satisfying the hypotheses of Theorem 4.6 on a manifold $V = M \times R$, with (M, e) u.l. Sobolev regular, has a solution u on $V_\tau := M \times [0, \tau]$ with given Cauchy $(\varphi, \psi) \in H_{s,u.l} \times H_{s-1,u.l}$ on M_0, if $s > \frac{n}{2} + 2$ and τ is small enough. The solution has a restriction to each M_t, $t \in [0, \tau]$, such that*

$$u(., t) \in H_{s,u.l}, \quad \partial_t u(., t) \in H_{s-1,u.l}.$$

If the system is hyperquasilinear, the theorem holds for $s > \frac{n}{2} + 1$.

5 Global existence

We commented in Section 3 on the global existence of solutions of the Cauchy problem for linear systems; it holds for regularly sliced Lorentzian manifolds. For the quasilinear wave equations that we have been considering a global existence theorem results from the energy estimates if the norms $\tilde{H}_s \times H_{s-1}$ of $u(t,.)$, $\partial_t u(t,.)$ of a solution u defined for $t \in [0,\ell)$ can be *a priori* bounded by a function of t, continuous for $t \in [t_0,\infty)$. Indeed, suppose we have proved that any solution with data in $\tilde{H}_s \times H_{s-1}$ at time $t=0$ is such that the $\tilde{H}_s \times H_{s-1}$ of $u(t,.)$, $\partial_t u(t,.)$ is *a priori* bounded by a continuous function of t, for $t \in [0,\infty)$, that is u does not blow up in a finite time; then an elementary reasoning shows that the solution exists globally on V. Let ℓ_0 define the maximal interval where the solution u exists. Suppose ℓ_0 is finite. Then, by choosing ε small enough it is possible to construct a solution with initial data induced on $t=\ell_0-\varepsilon$ by the solution u and which exists up to some $t>\ell_0$, contradicting the maximality of the interval of existence of u.

We treat some examples, which indicate applications and difficulties.

5.1 *Semilinear systems*

5.1.1 *The equation* $\Box_{\mathbf{g}} u - au^3 = 0$, $n=3$

The equation $\Box_\eta u - u^3 = 0$ on Minkowski spacetime M^4 was the first equation for which was proved[35] the global existence of solutions, without restriction on the size of initial data. The same result was proved[36] for the equation

$$\Box_{\mathbf{g}} u - au^3 = 0 \tag{5.1}$$

on a regularly sliced (hence globally hyperbolic) $3+1$ manifold by using energy estimates, as follows.

Section 4.2 has shown the local in time existence of a solution of the Cauchy problem for initial data in $H_1 \times H_0$, if $a \in C^0(T)$. To show global existence, in the case $a \geq 0$, we introduce a modified stress energy tensor defined by

$$T_{\alpha\beta} \equiv \partial_\alpha u \partial_\beta u - \frac{1}{2} g_{\alpha\beta} \partial_\lambda u \partial^\lambda u - \frac{1}{4} a g_{\alpha\beta} u^4 \tag{5.2}$$

Lemma 5.1 *When u is a solution of Equation (5.1) the $\mathbf{g}$-divergence of the stress energy tensor vanishes.*

$$\nabla_\alpha T^{\alpha\beta} = 0. \tag{5.3}$$

Proof Computation shows that

$$\nabla_\alpha T^{\alpha\beta} \equiv (\nabla_\alpha \partial^\alpha u - au^3)\partial^\beta u.$$

□

[35] Segal, I. E. (1963) *Ann. Math.*, **78**(2), 339–64, used energy estimates and semi-groups.
[36] Choquet-Bruhat, Y. and Cagnac F. (1984) *J. Math. Pures Appl.*, **63**, 377–90.

Theorem 5.2 *Equation (5.1) has, on a regularly sliced Lorentzian manifold, a global solution taking initial data* $u(0,.)$, $\partial_t u(0,.) \in H_1 \times H_0$.

Proof Using the definition of the energy momentum vector $\mathcal{P}$ associated to a stress energy tensor and a vector field X, $\mathcal{P}^\alpha := T^{\alpha\beta}X_\beta$, we have

$$\mathcal{P} \equiv \partial u(X.\partial u) - \frac{1}{2}X(\partial u.\partial u) - \frac{1}{4}au^4 X, \tag{5.4}$$

therefore, using previous results, $\mathcal{P}$ is timelike, with opposite orientation to X, if X is timelike.

We set, with the notation of Section 2

$$y_{X,t}(u) \equiv \int_{M_t} \left\{ G_{X,n}(\partial u, \partial u) + \frac{1}{4}au^4 \right\} \mu_g \tag{5.5}$$

Under the hypothesis made on $\mathbf{g}$ and if $a \geq 0$ we have

$$|T(u(t,.))| \leq C y_{X,t}(u). \tag{5.6}$$

The integration of the divergence of $\mathcal{P}$ gives that a solution of (5.1) satisfies an integral inequality

$$y_{X,t}(u) \leq y_{X,0}(u) + C_1 \int_0^t y_{X,t}(u) d\tau, \tag{5.7}$$

with

$$C_1 := \frac{C}{2} \underset{V}{Sup} N |\mathcal{L}_X \mathbf{g}|. \tag{5.8}$$

The Gronwall lemma applied to the linear integral inequality satisfied by $y_{X,t}(u)$ shows that this function possibly has exponential growth, but does not blow up in a finite time. The solution u extends therefore to a global in time solution. □

5.1.2 *Wave maps in dimension* $1+1$

The good properties of the Sobolev embedding and multiplication properties in dimension $n = 1$ permit global existence proofs for semilinear wave equations in two-dimensional spacetimes which do not extend to higher dimensions. The boundedness of the first and second energies can be proved for wave maps with source a regularly sliced two-dimensional Lorentzian manifold and target a complete Riemannian manifold[37]; this boundedness is sufficient to prove the global existence of these wave maps, even for large initial data. For higher dimensions the general problem is still unsolved, in spite of partial results[38].

[37] Choquet-Bruhat, Y. (1987) *Ann. IHP.*, **46**(1), 97–111. See previous work on wave maps from two-dimensional Minkowski spacetime with the method of characteristics by Gu Chao Hao *loc. cit.*

[38] See Chapters 15 and 16, and unpublished results of Terence Tao in the case $n = 2$ for large data.

5.1.3 *Yang–Mills equations in dimension* $3+1$. *Kirchhoff–Sobolev formula*

The Yang–Mills equations are not an ordinary system of semilinear wave equations. Global existence on Minkowski spacetime, for large data, has been proved by Eardley and Moncrief[39] using the physical Yang–Mills energy, but also a L^∞ bound obtained by an integral equation on the light cone; they used several different gauges, including the **Cronström gauge** for their light cone estimate, and the estimate of the Yang–Mills potential in terms of the fields. The results have been extended to Lorentzian manifolds by Chruściel and Shatah[40].

The integral equation is obtained, in the case of Minkowski spacetime M^4, by the Kirchhoff formula which gives the value at a point of the solution of a Cauchy problem for an inhomogeneous wave equation

$$\Box_\eta u = f$$

in terms of the integral of f on the cone and integrals on the initial manifold of a combination of the Cauchy data. An analogue of the Kirchhoff formula, but now an integral equation on the characteristic cone, baptized by Klainerman and Rodnianski[41] the **Kirchhoff–Sobolev formula**, was obtained by Sobolev[42] and used[43] to prove local in time existence of solutions of the Cauchy problem for quasilinear systems of second-order equations, in particular Einstein equations in wave gauge. Recently Klainerman and Rodnianski have written the integral on the cone putting in evidence geometric elements and lowering the assumed regularity of the metric, in view of proving the L^2 **bounded curvature conjecture**, local in time existence of a solution of Einstein equations with Cauchy data in $H_{2,loc} \times H_{1,loc}$, after applying the formula to the wave equation satisfied by the Riemann tensor (see Chapter 8). Moncrief[44], inspired by the analogy between the wave equations satisfied by the Riemann tensor and by the Yang–Mills field, had been working on the same project and also wrote an integral equation for the Riemann tensor, but with the Hadamard formulation of the elementary solution of the wave equation. Moncrief found formulae, in some sense analogous to the relation between the Yang–Mills potential and field in Cronström gauge, which give the connection and the metric in terms of the Riemann tensor[45] in geodesic coordinates.

[39] Eardley, D. and Moncrief, V. (1982) *Commun. Math. Phys.*, **83**, 171–92 and 193–203; result generalized using only energy estimates in Klainerman, S. and Machedon, M. (1995) *M. Ann. Math.*, **142**, 39–119.

[40] Chruściel, P. and Shatah, J. (1997) *Asian J. Math.*, **1**, 530–48.

[41] Klainerman, S. and Rodnianski, I. (2007) *J. Hyp. Diff. Eq.*, **4**, 401–33.

[42] Sobolev, S. (1936) *Math. Sbornik*, **1**(N.S.), 39–71. See further properties in Choquet-Bruhat, Y. (1952) *Annali di Matemetica*, serie 4, **64**, 191–228.

[43] Fourès (Choquet-Bruhat), Y. (1952) *Acta Matematica*, **88**, 141–225.

[44] Moncrief, V. (2005), Newton Institute, preprint, Cambridge.

[45] Another formulation of such relations can be found in the book by Christodoulou and Kainerman; see Chapter XV.

5.2 *Quasilinear equations*

The proof of global existence of solutions of quasilinear wave equations via energy estimates is made more difficult by the presence of the unknowns in the coefficients defining the energies. A way out is to suppose some *a priori* bound of the unknowns, deducing then energy estimates and using what is called a bootstrap argument. See examples in Chapters 15 and 16, Section XVI.4.4. The method yields usually global existence only for small initial data.

APPENDIX IV

GENERAL HYPERBOLIC SYSTEMS

1 Introduction

Hyperbolic partial differential equations govern physical phenomena which propagate with a finite velocity. They play a great role in mathematical physics, particularly in relativistic theories, where the causality principle requires that the speed of light be a maximum for the propagation of all observables.

A system of partial differential equations is called a **hyperbolic system** if it admits a well-posed Cauchy problem with data in spaces where only a finite number of derivatives are involved (more precisely in Sobolev spaces H_s), and the solution of this Cauchy problem manifests a finite propagation speed. By this property we mean that the value of the solution at a point depends only on the value of the data on a compact subset of the initial manifold. The quasidiagonal quasilinear second-order partial differential systems studied in the previous appendix are hyperbolic systems. They are special cases of **Leray hyperbolic systems**[1]. These second-order systems can also be written, by introducing derivatives as new unknowns[2], as **first-order symmetric hyperbolic systems**[3]. In this appendix we give results concerning these two types of hyperbolic systems. Leray systems are much more general, but their criterion for hyperbolicity is not satisfied by some symmetric hyperbolic systems.

In Section 3 we give results concerning Leray–Ohya hyperbolic systems. Such systems are not hyperbolic in the sense given above – the solutions are in the so-called Gevrey classes of C^∞ functions – but the solutions enjoy a finite propagation speed, a fundamental property in Relativity. The Leray–Ohya theory has no analogue for symmetric first-order systems.

2 Leray hyperbolic systems

2.1 *Case of one equation*

2.1.1 *Linear equation*

Consider a linear differential operator of order m on an $n+1$-dimensional manifold V, acting on a scalar function. Such an operator reads in local coordinates

[1] Leray, J. (1953) *Hyperbolic Differential Equations*, mimeographed, IAS Princeton.

[2] This procedure does not give better results and obscures the physical and geometrical properties.

[3] Introduced by Friedrichs, O. (1954) *Commun. Pure Appl. Math.*, **7**, 345–92.

on an open set $U \subset R^{n+1}$

$$L := \sum_{|\alpha| \leq m} a^{\alpha} \partial_{\alpha}^{|\alpha|}, \tag{2.1}$$

where α is a multi-index and $|\alpha|$ the total order of the derivations it defines,

$$\alpha := \alpha_0 \alpha_1, \ldots \alpha_{m-1} \quad \text{and} \quad |\alpha| := \alpha_0 + \alpha_1 + \cdots + \alpha_{m-1}. \tag{2.2}$$

The coefficients a^{α} are given functions on U. The operator L is written in invariant form on V if the operator ∂ denotes a covariant derivative in a given smooth connection on V.

The principal part of L is the operator

$$\sum_{|\alpha|=m} a^{\alpha} \partial_{\alpha}^{|\alpha|}. \tag{2.3}$$

This principal part defines at each point $x \in V$ a cone $P_x^* = 0$ in the cotangent vector space $T_x^* V$ to V. This cone has vertex x and a polynomial equation in the components of a vector X

$$X \in T_x^* V, P_x^* := \sum_{|\alpha|=m} a^{\alpha}(x) X_{\alpha} = 0, \quad X_{\alpha} := X_{\alpha_0} \quad X_{\alpha_{m-1}}. \tag{2.4}$$

The polynomial P_x^* is said to be a **hyperbolic polynomial** if there exists a point $Y \in T_x^* V$ such that each straight line through Y which does not go through the vertex x, $X = 0$, intersects the cone $P_x^* = 0$ in m distinct points. It can be shown that the closure of the set of such points Y consists of two closed convex opposite cones with non-empty interiors, $C_x^{*,+}$ and $C_x^{*,-}$, whose boundaries belong to $P_x^* = 0$ (but not necessarily coinciding with it; see example).

The operator L is said to be **hyperbolic at** x if its principal part defines a hyperbolic polynomial.

With P_x^* is associated a first-order non-linear differential equation for a scalar function f, called the **eikonal** equation, obtained by replacing in 2.4 X_{α} by $\partial_{\alpha} f$,

$$F(x,p) := \sum_{|\alpha|=m} a^{\alpha} \partial_{\alpha_0} f \ldots \partial_{m-1} f = 0, \quad p := \partial f \tag{2.5}$$

The **characteristic tangent vectors** to V at x are defined to be the vectors in $T_x V$ with components $Y^{\alpha}(p) := \frac{\partial F}{\partial p_{\alpha}}(x,p)$ with p satisfying $F(x,p) = 0$, hence

$$Y^{\alpha} p_{\alpha} = 0. \tag{2.6}$$

If the manifold V is endowed with a given pseudo-Riemannian metric and if the tangent and cotangent space are associated via this metric, then the relation (2.6) expresses the orthogonality in this metric of Y and p.

The vectors $Y(p)$ satisfying (2.6) constitute a cone $P_x = 0$, the dual cone of the cone $P_x^* = 0$. The cone $P_x = 0$ is contained in the cone C_x, dual to the cone

C_x^*. We denotes by C_x^+ and C_x^- the convex cones in the tangent plane duals to $C_x^{*,+}$ and $C_x^{*,-}$.

Example 2.1 *Suppose that the cone $P_x^* = 0$ is the union of the second-order cone of a metric g of Lorentzian signature $(- + \cdots + +)$ and a hyperplane orthogonal in this metric to a vector u*

$$P_x^*(X) \equiv g_x^{\alpha\beta} u_x^\lambda X_\lambda X_\alpha X_\beta = 0. \tag{2.7}$$

The polynomial P_x^* is hyperbolic if the hyperplane $u_x(X) = 0$ is spacelike. If we associate the cotangent and tangent space via the metric g_x, then the second-order cone $g_x(X, X) = 0$ appears as its own dual, while the dual of the hyperplane is a straight line with support u_x. The cone C_x is the cone $g_x^{\alpha\beta} X_\alpha X_\beta \leq 0$ if the hyperplane is spacelike, i.e. if u_x is timelike.

In the above example, the manifold V is **time oriented** if it is possible to distinguish continuously on V the convex cones C_x^+ and C_x^-, called respectively future and past cones. The definition is extended by analogy to general hyperbolic equations on manifolds. The causal past, respectively causal future, of a set is defined as it is in spacetimes (see Chapter 12), but now with the cones C_x^- and C_x^+; that is,

Definition 2.2 *A path on V is said to be L-causal if its tangent at each point $x \in V$ is in C_x^+.*

Leray also proved that the boundary of the L-causal past (or future) is generated by rays associated with the operator L, with the definition:

Definition 2.3 *The **rays** associated with L are the projection on V of the bicharacteristics of the first-order eikonal equation (2.5). These bicharacteristics are the curves in the cotangent space[4] T^*V which satisfy the ordinary differential system (the parameter λ on the curve is called the canonical parameter)*

$$\frac{dx^\alpha}{\partial F/\partial p_\alpha} = -\frac{dp_\beta}{\partial F/\partial x^\beta} = d\lambda, \quad \textit{for all } \alpha \textit{ and } \beta. \tag{2.8}$$

In the case where the principal operator is the wave operator of a Lorentzian metric g, the rays (called in that case light rays, or null rays) are the null geodesics of the metric.

Leray was inspired[5] by the definition of hyperbolicity in an open set $U \subset R^{n+1}$, which required that when all cones $C_y^{*,+}$, $y \in U$, are translated so as to have their vertex at the same point x, then the resulting set of cones has a non-empty intersection. He introduced the following definition for an operator L hyperbolic

[4] A point on a bicharacteristic is a pair (x, p), $x \in V$, $p \in T_x V$.

[5] Leray, J. *Hyperbolic Differential Equations*, mimeographed, IAS Princeton; argument briefly reproduced in CB-DM1 VI C 6.

at each point $x \in V$, to be called **globally hyperbolic**. The global hyperbolicity considered in Chapter 12 is a particular case of Leray hyperbolocity.

Definition 2.4 *The hyperbolic differential operator L is **globally hyperbolic** on V if the set of L-causal paths between two points $x, y \in V$ is compact (in the topology of sets of paths; see Chapter 12).*

Leray (1953) has proved future [respectively past] global existence theorems for a solution of the Cauchy problem for the linear globally hyperbolic equation on a manifold V

$$Lu = f, \quad f \text{ a given function} \tag{2.9}$$

when the initial data are given on a submanifold M of codimension one which Leray called **compact towards the past** [respectively future] with the following definition[6].

Definition 2.5 *A subset of a manifold V on which is defined a globally hyperbolic operator L is called L-past compact if its intersection with the L-future of any compact subset of V is a compact set. A similar definition holds by exchanging past and future.*

Leray also obtained, for a general hyperbolic equation (hence also for hyperbolic systems; see the next section), the **domain of dependence property**: the solution at a point in the L-future of M depends only on the L-causal past of this point.

Leray's conditions on the nature of the initial manifolds, apart from its compactness towards the past (or future), are very general. He obtained his theorems in Sobolev spaces, or local Sobolev spaces, on the manifold V. Under appropriate hypotheses on the coefficients of L the solution of an hyperbolic equation of order m is smoother than f by $m-1$ order of derivation. The energy estimates used by Leray in his proofs led to some loss of regularity of the solution with respect to the Cauchy data.

The regularity of the coefficients required for Leray's existence theorem was lowered by his student Dionne[7] and the loss of regularity under evolution suppressed. But Dionne's study was limited to $V \equiv R^{n+1}$ and to the case of one equation. Moreover he studied only the classical Cauchy problem with initial data on R^n, assumed to be uniformly spacelike with respect to the L-hyperbolic structure. Nobody seems to have had the courage, or the motivation, to extend these improvements to the beautiful Leray results for general systems of arbitrary order on manifolds[8].

[6] Of course this definition applies to the manifold V, not to its compactification by a Penrose formalism or other.

[7] Dionne, P. (1962) *J. Anal. Math. Jerusalem*, **10**, 1–90.

[8] Leray's idea for solving a generalized Cauchy problem has been used by students of F. Cagnac in works on the characteristic Cauchy problem.

2.1.2 *Quasilinear equation*

A quasilinear equation of order m for a scalar unknown u on a manifold V reads

$$\sum_{|\alpha|=m} a^{\alpha}(u, \partial u, \ldots, \partial^{m-1}u)\partial^{m}_{\alpha}u + b(u, \partial u, \ldots, \partial^{m-1}u) = 0, \tag{2.10}$$

where ∂ denotes the covariant derivative in some given smooth connection on V. The equation is said to be hyperbolic for a specified u if it is hyperbolic when u is replaced by this function in the coefficients a^{α} and b.

The lowering on the hypothesis of regularity appearing in the Leray local existence and uniqueness theorem on a manifold has been written up by Dionne only in the case of R^{n+1}, and still using Sobolev spaces on V. However, we state the general theorem, which we believe to be true.

Theorem 2.6 *Let $V \equiv M \times R$, with (M, e) a Sobolev regular Riemannian manifold. Assume that the coefficients a and b of Equation (2.10) are smooth functions of $u, \partial u, \ldots, \partial^{m-1}u$, and that the operator L is globally hyperbolic on V whenever the pointwise values of $(u, \partial u, \ldots, \partial^{m-1}u)$ belong to some given open set*[9] *of R^N, and assume that the submanifolds M_t, $0 \leq t \leq t_1$ are then uniformly L-spacelike. Let*

$$\partial^{p}u|_{M} \in H^{u.loc}_{s-p}, \quad p = 0, 1, \ldots m-1, \quad s > \frac{n}{2} + m - 1, \tag{2.11}$$

be Cauchy data on M_0

Then there exists a number $T \leq t_1$ such that this Cauchy problem has one and only one solution $u \in E^{u.loc}_{s}(T)$.

The number T is the same[10] *for all $s \geq s_0$, s_0 the smallest integer such that $s_0 > \frac{n}{2} + m - 1$. The solution is C^{∞} if the Cauchy data are C^{∞}.*

Remark 2.7 *To have a simple statement we have not included direct dependence in x of the coefficients. We refer the reader to the quoted papers by Leray and Dionne for the relevant statements.*

Remark 2.8 *In General Relativity the gravitational field is always present and propagates at the speed of light. The relativistic causality principle imposes the requirement that all other physical fields propagate with a smaller or equal speed. In consequence, for the Einstein equations with classical sources, which are a system of differential equations for various fields on a manifold V, including a Lorentzian metric g, the cones C_x must coincide with the causal cones of the spacetime (V, g). Global hyperbolicity of a hyperbolic system for Einstein equations and their sources is the global hyperbolicity of the spacetime (V, g).*

[9] This formulation is correct only if V is diffeomorphic to R^{n+1}. The formulation for a general V being cumbersome, we leave to the reader the task of finding the correct formulation in applying the theorem to other manifolds, depending on the reader's particular interest.

[10] Choquet-Bruhat, Y. (1972) *J. Gen. Rel. Grav.*, **2**(4), 359–62.

2.2 *Leray hyperbolic systems*

A system of N differential equations for N scalar unknowns[11] u^A, $A = 1, \ldots, N$, on a differentiable manifold V is said to be quasilinear if it is possible[12] to attach to each unknown an index m_A and to each equation an index n_B (hence to find $2N$ integers m_A and n_A), such that the system reads (no summation over indices m and n)

$$F^A(u) \equiv a_B^{A,\alpha}(\tilde{\partial}^{m_C - n_A - 1} u^C)\partial_\alpha^{m_B - n_A} u^B + b^A(\tilde{\partial}^{m_C - n_A - 1} u^C) = 0 \qquad (2.12)$$

where $\tilde{\partial}^m$ denotes the set of derivations of order less or equal to m. The operator $a_B^{A,\alpha}(\tilde{\partial}^{m_C - n_A - 1} u^C)\partial_\alpha^{m_B - n_A} u^B$ is called the principal part relative to u^B in F^A; it may be zero if $m_B - n_A < 0$.

The **characteristic determinant** at $x \in V$ of the system (2.12) for a given u, is the polynomial P_x^*, of degree $\ell := \Sigma m_B - \Sigma n_A$, in the cotangent space to V at x with elements

$$p_B^A(X) := a_B^{A,\alpha}(\tilde{\partial}^{m_C - n_A - 1} u^C(x))X_\alpha, \quad |\alpha| = m_B - n_A. \qquad (2.13)$$

The system is said to be **Cauchy regular** if the characteristic determinant is not identically zero. Then, when u is known the equation $P_x^* = 0$ defines, as in the case of one equation, a cone in the cotangent plane and its dual a cone $P_x = 0$ in the tangent plane. The definitions of hyperbolicity at x is the same as for one equation. The definition of global hyperbolicity extends also for a given u. A system with these two properties is called a **Leray-hyperbolic system**. Leray has proven a local existence and uniqueness theorem for a solution of the Cauchy problem for such hyperbolic systems. The extension to systems of the results of Dionne do not appear in the literature, but can be proved by using **quasidiagonalisation**.

Definition 2.9 *The system (2.12) is called quasidiagonal if the principal terms are diagonal, i.e. the system reads*

$$F^A(u) \equiv a^\alpha(\tilde{\partial}^{\ell - 1} u)\partial_\alpha^\ell u^A + b^A(\tilde{\partial}^{\ell - 1} u) = 0 \qquad (2.14)$$

Theorem 2.10 *A Cauchy regular system can always be written in quasidiagonal form*[13].

Proof Denote by $\hat{p}_A^B$ the cofactor of the element p_B^A in the determinant P^*, polynomial in X of order $\ell - m_B + n_A$. It holds that

$$\hat{p}_A^C(X)p_B^A(X) = \delta_B^C P(X). \qquad (2.15)$$

[11] Which can be the representatives in a chart of a tensor or other geometric object.

[12] For explanations about the origin of this definition, and the uniqueness of the corresponding characteristic determinant under change of choice of indices m and n, see Choquet-Bruhat, Y. (1966) *J. Math. Pures Appl.*, **45**, 371–86.

[13] Choquet-Bruhat, Y. (1966) *J. Math. Pures Appl.*, **45**, 371–86.

Equation (2.15) does not hold as it stands if we replace the components X_α of X by the partial derivatives ∂_α, because the operator $p_A^C(\partial)$ has coefficients which depend on x. However, the operators on the left- and right-hand sides of (2.15) have the same principal parts, of order ℓ. Some difficulty appears with the coefficients which would contain derivatives of order greater than $\ell - 1$. It can be checked that this does not appear if

$$\ell \geq \underset{A}{Sup} m_A - \underset{B}{Inf} n_B, \tag{2.16}$$

in particular if all integers $m_A - n_B$ are ≥ 0; that is, if all unknowns appear in each equation. If it is not the case one can prove the theorem recursively. □

A consequence of the previous theorem is the following one.

Theorem 2.11 *Theorem 2.6 for a scalar equation of order ℓ extends to Leray hyperbolic systems with indices such that $\Sigma m_B - \Sigma n_A = \ell$.*

Leray has extended his result to the case of systems **quasidiagonal by blocks**, eliminating then some spurious multiplicities of characteristics.

Definition 2.12 *The system (2.12) is said to be quasi-diagonal by blocks if the non-zero elements of the characteristic determinant are in square blocks around the diagonal, as indicated in the example below (letters denoting differential operators)*

$$\begin{pmatrix} a & b & & & & \\ c & d & 0 & & 0 & \\ & & & & & \\ 0 & & e & & 0 & \\ & & & f & g & h \\ & & & & & \\ 0 & & 0 & i & j & k \\ & & & l & m & n \end{pmatrix} \tag{2.17}$$

Theorem 2.13 *A system quasidiagonal by blocks is Leray-hyperbolic at x if each block is Leray-hyperbolic, and the intersection of the cones $C_x^{*,+}$ has a non-empty interior. The dual of this intersection defines the causal cone of the system.*

Remark 2.14 *A system quasidiagonal by blocks can be written as a system where each block is quasidiagonal. Such a system reads (no summation on A)*

$$F^A(u) \equiv a^{A,\alpha}(\tilde{\partial}^{m_C - n_A - 1} u^C)\partial_\alpha^{m_A - n_A} u^A + b^A(\tilde{\partial}^{m_C - n_A - 1} u^C) = 0. \tag{2.18}$$

Theorem 2.15 *The Leray existence and uniqueness theorem extends to Leray-hyperbolic systems.*

We leave the precise, Dionne-like, formulation and its proof to the reader.

Remark 2.16 *It may be possible*[14] *to diagonalize a system by using quotients of cofactors by a common polynomial, thus making disappear further spurious multiplicity of characteristics, and proving the hyperbolicity of systems.*

3 Leray–Ohya hyperbolic systems

The **Leray–Ohya hyperbolic** systems, studied by Leray and Ohya[15], were called by them "**hyperboliques non strictes**". Such quasilinear systems are of the form (2.12) and their characteristic polynomial P_x^* still defines two opposite convex cones $C_x^{*,+}$ and $C_x^{*,-}$ in the cotangent space, and the points where a straight line passing through a point in C_x^* cuts $P_x^* = 0$ are all real points, but they may be not all distinct: the characteristics may be multiple. The solutions of such systems enjoy the domain of dependence property if they are defined as follows.

Definition 3.1 *The quasilinear system (2.14) is called Leray–Ohya hyperbolic for a specified u if (we omit writing x)*

1. *The characteristic polynomial P^* is a product of hyperbolic polynomials*
$$P^*(X) = P_1^*(X) \dots P_N^*(X). \tag{3.1}$$
2. *The cones $P_i^*(X) = 0$ all contain in their interior the same cone $C_x^*(X) = 0$.*
3. $$Sup \ell_i \geq \underset{A}{Sup} m_A - \underset{B}{Inf} n_B, \tag{3.2}$$
where ℓ_i is the degree of the polynomial P_i^*.

The number
$$\sigma = \frac{N}{N-1} \tag{3.2}$$
is called the **Gevrey index** of the system. This index is $\sigma = \infty$ if the system is hyperbolic ($N = 1$). Solutions of Leray–Ohya hyperbolic systems are obtained in **Gevrey classes**. The definition of Gevrey classes appear in the literature only on open sets of R^n. They are classes of C^∞ functions whose successive derivatives satisfy inequalities, but inequalities weaker than those satisfied by analytic functions which give the convergence of Taylor series. To a function in a Gevrey class there corresponds only a formal series. Their virtue is that they are not determined by their value in an open set. We formulate definitions and results in the case of an open set $U \subset R^n$. There is no conceptual difficulty in generalizing the formulation to manifolds, which are locally diffeomorphic to R^n, but it has not been written up before, and it would be too heavy for this appendix.

[14] Choquet-Bruhat (1966) *loc cit.* See applications in Chapter 9.
[15] Leray, J. and Ohya, Y. (1966) *Math. Annalen*, **162**, 228–36.

Definition 3.2 *A C^∞ function f belongs to the Gevrey class $\gamma_2^{(\sigma)}(U)$, $1 \le \sigma < \infty$, if*

$$\underset{\alpha}{Sup} \frac{1}{[1+|\alpha|]^\sigma} ||\partial^\alpha f||_{L^2(U)} < \infty. \tag{3.3}$$

It can be shown that functions in $\gamma_2^{(1)}(U)$ are analytic functions, while functions in $\gamma_2^{(\sigma)}(U)$, $\sigma > 1$, with compact support in U may be not identically zero in U.

Leray and Ohya prove an existence and uniqueness theorem of a solution of the Cauchy problem with initial data in a Gevrey class $\gamma_2^{(\sigma)}$ for quasilinear equations, or systems, of Gevrey index σ, which they called "hyperboliques non strictes". These solutions enjoy the domain of dependence property.

4 First-order symmetric hyperbolic systems

4.1 *FOSH systems on R^{n+1}*

4.1.1 *Definition*

Classically, one considers a system of N first-order linear partial differential equations on $R^n \times R$ for a set $u = (u_I, I = 1, \ldots, N)$ of N functions. It reads

$$Lu := \mathcal{M}^\alpha \frac{\partial}{\partial x^\alpha} u + \mathcal{A} u = f, \quad \text{i.e.} \quad \mathcal{M}^{\alpha IJ} \frac{\partial}{\partial x^\alpha} u_I + \mathcal{A}^{IJ} u_I = f^J \tag{4.1}$$

where $f = (f^I, I = 1, \ldots, N)$ is a set of given functions, $\mathcal{M}^\alpha$ and $\mathcal{A}$ are given $N \times N$ matrices.

We recall that the $C^0(M_t)$ norm of any matrix $\mathcal{A}$ is defined by:

$$|\mathcal{A}|_{C^0(M_T)} := \underset{M_t, |X|=1}{Sup} |\mathcal{A}^{IJ} X_J X_I|, \quad |X|^2 = \sum_I |X_I|^2. \tag{4.2}$$

Definition 4.1 *The system (3.1) is called symmetric if the matrices $\mathcal{M}^\alpha$ are symmetric.*

A matrix $\mathcal{M}^t$ is said to be continuously positive definite if there exists a positive and continuous function $C_H(t) > 0$ such that

$$\underset{M_t, |X|=1}{Inf} \mathcal{M}^{t,IJ} X_J X_I \ge C_H(t), \quad |X|^2 = \sum_I |X_I|^2. \tag{4.3}$$

The following definition will be justified by a theorem to come.

Definition 4.2 *A first-order symmetric differential system is called hyperbolic (FOSH) for the time function $x^0 = t$ if the quadratic form defined by the matrix $\mathcal{M}^t$, coefficient of $\frac{\partial}{\partial t}$, is continuously positive definite.*

4.1.2 *Energy estimate*

To a FOSH system is associated an **energy functional** defined for each t by

$$E_t(u) := \int_{M_t} \mathcal{M}^t(u,u) d^n x, \tag{4.4}$$

where $d^n x$ is the volume element of the Euclidean space R^n and $M_t := R^n \times \{t\}$)

Under the hypothesis that $\mathcal{M}^t$ is continuously positive definite the energy is bounded below by the $L^2(M_t) \equiv L^2(R^n)$ norm of the t-dependent function u,

$$E_t(u) \geq C_H(t) ||u||_{L^2(M_t)}. \tag{4.5}$$

Theorem 4.3 *Suppose that the matrices $\mathcal{M}^\alpha$ of a FOSH system and belong to $C^1(R^n \times [t_0, T])$, $\mathcal{A} \in C^1(R^n \times [t_0, T])$ and $f \in L^1([t_0, T), L^2(R^n))$. Then any solution u in $C_0^1([t_0, T) \times R^n)$ is such that:*

$$E_t^{\frac{1}{2}}(u) \leq \left\{ E_{t_0}^{\frac{1}{2}}(u) + C_1(t) \int_{t_0}^t ||f||_{L^2(R^n)} dt \right\} \exp \frac{1}{2} \int_{t_0}^t C_2(t) dt. \tag{4.6}$$

where $C_1(t)$ and $C_2(t)$ are positive and continuous functions of t. The function C_1 depends on C_H and on the C^0 norms of the $\mathcal{M}^\alpha$. The function C_2 depends on the C^0 norms of $\mathcal{A}$ and on the divergence of the vector valued matrix $\mathcal{M}^\alpha$.

Proof The contracted product of (4.1) with u gives, if the matrices M^α are symmetric,

$$\frac{1}{2} \frac{\partial}{\partial x^\alpha} (\mathcal{M}^{\alpha IJ} u_J u_I) - \frac{1}{2} \left(\frac{\partial}{\partial x^\alpha} \mathcal{M}^{\alpha IJ} \right) u_J u_I + \mathcal{A}^{IJ} u_J u_I - f^J u_J = 0. \tag{4.7}$$

Under the hypothesis made on $\mathcal{M}$, and if[16] $u \in C_0^1([t_0, T) \times R^n)$, integration on the strip $R^n \times [t_0, T]$ and the Stokes formula lead to the **energy equality**

$$E_T(u) = E_{t_0}(u) + \int_{t_0}^T \left\{ \frac{1}{2} \left(\frac{\partial}{\partial x^\alpha} \mathcal{M}^{\alpha IJ} \right) u_J u_I - \mathcal{A}^{IJ} u_J u_I + f^J u_J \right\} d^n x dt. \tag{4.8}$$

Under the hypothesis made on $\mathcal{M}^t$ and $\mathcal{A}$ this equality implies the following **energy inequality**

$$E_T(u) \leq E_{t_0}(u) + \int_{t_0}^T \{ C_2(t) E_t(u) + ||f||_{L^2(M_t)} [C_H^{-1} E_t(u)]^{\frac{1}{2}} \} dt.$$

From the energy inequality one deduces (Gromwal lemma) the **energy estimate** (4.6). □

One obtains higher order energy estimates by differentiating the equation.

[16] C_0^k denotes C^k functions on $R^n \times R^{n+1}$ with trace in R^n of compact support.

4.1.3 *Existence and uniqueness theorem*

One deduces from energy estimates an existence and uniqueness theorem for a solution of the Cauchy problem for linear equations, by methods analogous to the ones exposed in detail for second order systems. A local existence theorem in the case of quasilinear equations is obtained by iteration and functional analysis theorems. Details of these proofs can be found in the books by Majda or by Dafermos[17]. The theorem is as follows.

Theorem 4.4 *Let*[18]

$$\mathcal{M}^\alpha(u)\frac{\partial}{\partial x^\alpha}u + f(u) = 0, \quad x^0 = t, \tag{4.9}$$

be a first-order symmetric system of partial differential equations on R^{n+1} *for a set of* N *scalar functions* u*. Assume that* $\mathcal{M}^\alpha$ *and* f *depend smoothly on* $u \in \Omega$*, some open set of* R^N*, and are such that* $\mathcal{M}^0 \in \bar{C}^0([0,T] \times R^n)$ *while* $\partial\mathcal{M}^\alpha$ *and* f *are in* $C^0([0,T), H_{s-1}(R^n))$ *if they depend directly on* t *and* x*, with* $s > \frac{n}{2} + 1$*. Assume* $\mathcal{M}^t$ *continuously positive definite in these domains. Then there is an interval* $[0,T') \subset [0,T)$ *such that the Cauchy problem*

$$u(0,.) = u_0, \quad u_0 \in H_s$$

has a unique solution $u \in E_s(T') := C^0([0,T'), H_s(R^n)) \cap C^1([0,T'), H_{s-1}(R^n))$*. if* $s > \frac{n}{2} + 1$*.*

The results extend readily to **symmetrizable** hyperbolic systems, i.e. systems such that there exists an $N \times N$ invertible matrix $\mathcal{S}$ such that the matrices $\tilde{\mathcal{M}}^\alpha := \mathcal{S}\mathcal{M}^\alpha$, with elements $\mathcal{S}^H_I\mathcal{M}^{IJ\alpha}$, are symmetric, with $\tilde{\mathcal{M}}^0$ positive definite.

Remark 4.5 *A first-order differential system is symmetrizable if the matrix* $\mathcal{M}^i(u)\xi_i$ *admits* N *real distinct eigenvalues with respect to* $\mathcal{M}^t(u)$*, supposed to be positive definite. The system is then in fact also a (strictly) hyperbolic Leray system. In the case of an eigenvalue of multiplicity* p*, the symmetric hyperbolicity follows from the existence of a* p*-dimensional eigenspace only in the case* $n = 1$ *unless one introduces pseudodifferential (non-local) operators which lead to another definition of hyperbolicity, called in recent literature* ***strong hyperbolicity*** *(see e.g. Nagy, G., Ortiz, O., and Reula, O., gr-qc 040123).*

4.2 *FOSH systems on a sliced manifold*

The extension of the above considerations to sliced hyperbolic manifolds is straightforward. However, first-order relativistic systems often do not appear as symmetric in natural coordinates. It is often useful to introduce other reference frames; however, one must be careful that energy estimates are obtained for energies related to the level sets of a time function $x^0 = t$; the matrix $\mathcal{M}^t$

[17] Dafermos, C. (1994) Lecture Notes in Mathematics 1640 (ed. T. Ruggeri), Springer.
[18] M^α and f can also depend directly on the spacetime (here R^{n+1}) point.

whose positivity defines hyperbolicity is the matrix of the coefficients of $\frac{\partial}{\partial t}$, not the matrix $\mathcal{M}^0$ of the coefficients of a Pfaff derivative ∂_0.

4.2.1 *First-order linear symmetrizable systems*

We consider a linear system of partial differential equations on a pseudo-Riemannian manifold (V, g) with unknown a tensor field, or a finite set of tensor fields u on V. Such a system reads, in an arbitrary moving frame (θ^α) where the components of u are u^I, $I = 1, \ldots, N$ and f is a given tensor field:

$$\mathcal{M}\nabla u = f, \quad \text{i.e.} \quad \mathcal{M}^\alpha \nabla_\alpha u = f \quad \text{i.e.} \quad \mathcal{M}^{IJ,\alpha} \nabla_\alpha u_J = f^I. \tag{4.10}$$

The expressions of the matrices $\mathcal{M}^\alpha$ and of the covariant derivative ∇_α depend on the chosen frame.

If $V = R^{n+1}$ the system looks in local coordinates like the one considered in the previous section; the same results apply, under the same hypothesis.

In relativistic theories one wants to treat more general manifolds and formulate theorems in a geometric form.

4.2.2 *Cauchy adapted frame*

We consider a sliced Lorentzian manifold $(V = M \times R, g)$ with metric in a Cauchy adapted frame

$$g = -N^2 dt^2 + g_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt). \tag{4.11}$$

We denote by M_t the spacelike slice $M \times \{t\}$, $t \in R$, by $\bar{g}$ the Riemannian metric $\bar{g}_t$ with components g_{ij} induced by g on M_t, and by $\bar{\nabla}$ the connection of $\bar{g}$.

A tensor field u on V can be decomposed into a finite set of t-dependent tensor fields v on M. A first-order linear differential system of the form (4.10) is equivalent to a system acting on such tensor fields which reads[19]

$$\mathcal{M}^{IJ,0} N^{-1} \bar{\partial}_0 u_I + \mathcal{M}^{IJ,h} \bar{\nabla}_h u_I + L^{IJ} u_I = f^J \tag{4.12}$$

where the f^J are components of given tensor fields. The coefficients L^{IJ} come from the difference of the covariant derivatives ∇_i and $\bar{\nabla}_i$. We assume that the system is symmetrized.

Definition 4.6 *The symmetric system (4.12) is hyperbolic for the time function t (**FOSH**, First-Order Symmetric Hyperbolic) if the matrix $\mathcal{M}^t$ of the coefficients of $\frac{\partial}{\partial t}$ on its left-hand side is continuously positive definite.*

Lemma 4.7 *A symmetric system is FOSH if the matrix $\mathcal{M}^0$ in a Cauchy adapted frame is positive definite.*

Proof The relations $\partial_0 = \frac{\partial}{\partial t} - \beta^i \frac{\partial}{\partial x^i}$ and $\partial_i = \frac{\partial}{\partial x^i}$ show that the matrices $\mathcal{M}^0$ and $\mathcal{M}^t$ are the same. □

We extend as follows the definition given by Equation (4.4) on R^{n+1}.

[19] For the definition of $\bar{\partial}_0$ and the relation between ∇_i and $\bar{\nabla}_i$, see chapter VI, local Cauchy problem.

Definition 4.8 *Suppose the matrix $\mathcal{M}^0$ is continuously positive definite. The energy at time t of a solution u of (4.2) is the positive integral*

$$E_t(u) := \frac{1}{2}\int_{M_t} \mathcal{M}^{HI,0} v_H v_I \bar{\mu}_t \tag{4.13}$$

One defines the hyperbolicity constant $C_H > 0$ by the same formula (4.5) as in Section 4.1.

We recall the following lemma, proved in Chapter 8, Equation (VIII.9.12).

Lemma 4.9 *For an arbitrary spacetime function F it holds that*

$$\int_{M_t} \partial_0 F \bar{\mu}_t = \frac{d}{dt}\int_{M_t} F\bar{\mu}_t + \int_{M_t} N\tau F \bar{\mu}_t, \tag{4.14}$$

where τ denotes the mean extrinsic curvature $tr_{\bar{g}}K$ of the slices M_t and $\bar{\mu}_t \equiv \mu_{\bar{g}_t}$ its volume element.

Theorem 4.10 *Let $u \in C_0^1$ be a solution on a strip $\Sigma_T = M \times [t_0, T]$ of the symmetric hyperbolic system (4.12), with time function t. Assume that*

1. *The energy $E_t(u)$ satisfies the equality*

$$E_T(u) = E_{t_0}(u) + \int_{t_0}^{T}\int_{M_t} \{\mathcal{Q}^{HI} u_H u_I - u_J f^J\}\bar{\mu}_t dt, \tag{4.15}$$

with

$$\mathcal{Q}^{HI} := -N\tau \mathcal{M}^{HI,0} + \bar{\partial}_0 \mathcal{M}^{HI,0} + \bar{\nabla}_h(N\mathcal{M}^{HI,h}) + \mathcal{A}^{HI}. \tag{4.16}$$

2. *Suppose that on each M_t the quadratic form $\mathcal{Q}$ is uniformly bounded through the quadratic form $\mathcal{M}^0$ and that f is square integrable; more precisely suppose that there exist continuous and bounded functions of $t \in [t_0, T]$ such that*

$$|\mathcal{Q}(u,u)| \leq C_2(t)\mathcal{M}^0(u,u), \quad \textit{i.e.} \quad |\mathcal{Q}^{HI}u_H u_I| \leq C_2(t)\mathcal{M}^{HI,0}u_H u_I. \tag{4.17}$$

Then the energy $E_t(u)$ satisfies the inequality

$$E_T(u) \leq E_{t_0}(u) + \int_{t_0}^{T}\{C_2(t)E_t(u) + C_2(t)E_t^{\frac{1}{2}}(u)\}dt. \tag{4.18}$$

Proof

1. Contracting the left-hand side of (4.12) with $\tilde{u}_J := u_H B_J^H$ and integrating this equation on Σ_T with respect to the spacetime volume element $N\bar{\mu}_t dt$ gives:

$$\int_{t_0}^{T}\int_{M_t} \{u_H \mathcal{M}^{HI,0}\bar{\partial}_0 u_I + u_H N\mathcal{M}^{HI,h}\bar{\nabla}_h u_I + N\mathcal{A}^{IJ}u_I u_J + N u_H f^J\}\bar{\mu}_t dt = 0. \tag{4.19}$$

Using the symmetry of the matrices $\mathcal{M}^\alpha$, we see that this equation may be written

$$\int_{t_0}^{T}\int_{M_t}\left\{\frac{1}{2}\partial_0(u_H\mathcal{M}^{HI,0}u_I)+\frac{1}{2}\bar{\nabla}_h(Nu_H\mathcal{M}^{HI,h}u_I)\right.$$
$$\left.-[\bar{\partial}_0\mathcal{M}^{HI,0}+\bar{\nabla}_h(N\mathcal{M}^{HI,h})]u_Hu_I+N\mathcal{A}^{IJ}u_Iu_J+Nu_Hf^J\right\}\bar{\mu}_t dt=0. \tag{4.20}$$

The use of the lemma and the Stokes formula give the announced result.

2. A straightforward consequence of the hypothesis. □

Higher order energy estimates can be obtained by differentiating the equations, and existence theorems by methods analogous to those given for the space R^{n+1}. We leave their formulation to the reader.

APPENDIX V

CAUCHY–KOVALEVSKI AND FUCHS THEOREMS

1 Introduction

The Cauchy–Kovalevski and Fuchs theorems are both existence theorems for analytic solutions of a Cauchy problem for systems of partial differential differential equations. They require the extension of the problem to holomorphic (i.e. complex analytic) functions[1].

The Fuchs theorem considers a singular Cauchy problem for particular systems called Fuchsian. The Cauchy–Kovalevski theorem is a particular case of the Fuchs theorem, but it has a much simpler proof. We recall this proof in the next section.

2 Cauchy–Kovalevski theorem

2.1 *Linear system*

We treat the case of a general linear first-order[2] system for a t-dependent tensor field u over an analytic manifold M,

$$\partial_t u = aD_x u + bu + f, \tag{2.1}$$

where f is given in the space $\mathcal{T}$ of tensor fields of the same type as u, b is a linear mapping from $\mathcal{T}$ into $\mathcal{T}$, D_x is the covariant derivative in some given analytic Riemannian metric e on M, and a is a linear map from $T^*M \otimes \mathcal{T}$ into $\mathcal{T}$, with T^*M the cotangent space to M. In local coordinates x^i, $i = 1, \ldots, n$ on M, a, b and f are represented by sets of scalar functions $a^{i,I}_J$, b^I_J, f^I, $I, J = 1, \ldots, N$, functions of t and the $x^{i'}s$ with the usual summations to define the operators on the representatives u^I and $D_i u^I$.

The Cauchy data is a tensor field $\phi \in \mathcal{T}$, represented by ϕ^I, $I = 1, \ldots, N$, functions of x^i.

We suppose that the Riemannian manifold (M, e) admits a holomorphic (i.e. complex analytic) extension $(\hat{M}, \hat{e})$ of positive width; that is, $\hat{M} \equiv M \times D$ with D a polydisc in R^n. The domains of the charts on $\hat{M}$ are of the form $\hat{\omega} \equiv \omega \times D$, with ω the domain of a chart in M. The local coordinates in $\hat{\omega}$ are the complex

[1] A limit of holomorphic functions in a fixed domain of C^n is a holomorphic function, while the limit of functions real analytic in a domain of R^n may not be analytic.

[2] Higher order systems can always be written in first-order form by the introduction of derivatives as new unknowns, but in the form 2.1 only if the manifolds $t =$ constant are not characteristic.

numbers $z^i \equiv x^i + iy^i$ with x^i local coordinates in ω and y^i coordinates in the polydisc; that is $|y^i| < d$, a positive number, $i = 1, \dots, n$.

We assume that the coefficients a, b, f admit a holomorphic extension to $\hat{M} \times \hat{R}$, with $\hat{R} = R \times (-T, T)$, $T > 0$; the data ϕ admits a holomorphic extension to $\hat{M}$.

Theorem 2.1 *(Local in space and time). Let x_0 be a point of M. There exists a domain $\Omega \subset M \times R$ defined by*

$$d(x, x_0) < \rho, \quad |t| < \tau, \quad \rho > 0, \quad \tau > 0 \tag{2.2}$$

such that the system has a solution u analytic in Ω and equal to ϕ on $\Omega \cap \{t = 0\}$.

The numbers ρ and τ depend only on $\hat{M} \times \hat{R}$ and on the maximum pointwise norm of a. They do not depend on f or ϕ.

The solution is unique.

Proof We reduce the problem to the case of zero Cauchy data by replacing the unknown u by another tensor w defined by

$$w(t, x) = u(t, x) - \phi(x). \tag{2.3}$$

The field w satisfies an equation of the same type as u, with the same a and b but with a different f, namely:

$$\partial_t w = aD_x w + bw + \tilde{f}, \quad \tilde{f} \equiv f + aD_x \phi + b\phi. \tag{2.4}$$

Consider a relatively compact domain ω of R^n included in a coordinate chart at x_0, with local coordinates such that $x_0^i = 0$ to simplify the writing. By the hypothesis the representatives of the metric e and of the coefficients $a, b, \tilde{f}$ extend to holomorphic functions in a polydisc of C^{n+1}

$$\hat{\Omega} := |z^i| < \rho, i = 1, \dots n, \qquad |z^0| < |\tau| \tag{2.5}$$

where $\rho > 0$ and $\tau > 0$ depend only on $\hat{M}$ and $\hat{R}$. Denote by H any one of these holomorphic functions; this is equal in the polydisc to the sum of its Taylor expansion at $(0, \dots 0)$[3],

$$H = \sum_{k_0, \dots k_n} H_{k_0 \dots k_n} (z^0)^{k_0} (z^1)^{k_1} \dots (z^n)^{k_n}, \quad \text{if} \quad |z^i| < \rho, \ |z^0| < \tau. \tag{2.6}$$

Therefore if the function H is bounded in $\hat{\Omega}$ it holds that

$$|H_{k_0 \dots k_n}| \leq \frac{C_H}{\tau^{k_0} \dots \rho^{k_n}} \quad \text{with} \quad C_H := \underset{\hat{\Omega}}{Sup} |H|. \tag{2.7}$$

The representatives of the successive derivatives at $x_0 = (0, \dots, 0)$ of a solution (if it exists) are computed from the expansions in $\hat{\Omega}$ of the Cauchy data and the

[3] Recall that a real analytic function is equal to the sum of its Taylor expansion in the neighbourhood of any point, but the size of this neighbourhood depends on the width of its holomorphic extension.

coefficients of the equation. If the obtained Taylor series converges the identification shows that it is a holomorphic solution of the problem and the only one. The coefficients of the Taylor series so computed for the representative of u in $\hat{\Omega}$ are smaller than coefficients of the series obtained by replacing representatives of a, b and $\tilde{f}$ by majorants; that is, quantities admitting Taylor expansions such that each coefficient is greater than the supremum of the absolute value of the corresponding coefficient of the Taylor expansion of the majorated quantity.

The function

$$H_1 := \frac{C_H}{1 - \{\tau^{-1}z^0 + \rho^{-1}(z^1 + \cdots + z^n)\}} \tag{2.8}$$

is holomorphic in the domain $\hat{\Omega}_1 := \hat{\Omega} \cap \{|\tau^{-1}z^0 + \rho^{-1}(z^1 + \cdots + z^n)| < 1\}$; this function H_1 is seen, by elementary calculus, to be a majorant of H. If we replace the coefficients of the given equation by such majorizing functions we obtain, for an unknown denoted w_1, an equation which reads in $\hat{\Omega}_1$

$$\partial_{z^0} w_1 = a_1 D_Z w_1 + b_1 w_1 + \tilde{f}_1. \tag{2.9}$$

We replace a_1, b_1, $\tilde{f}_1$ by their values[4], we set $\xi = \tau^{-1}z^0 + \rho^{-1}(z^1 + \cdots + z^n)$. The equation becomes, after multiplication by $1 - \xi$,

$$\alpha(\xi)\partial_\xi w_1 = C_b w_1 + C_{\tilde{f}}, \quad \text{with} \quad \alpha(\xi) := \tau^{-1}(1 - \xi) - \rho^{-1}C_a. \tag{2.10}$$

This linear differential equation admits a holomorphic solution with initial value zero in the domain where α^{-1} is a holomorphic function of ξ; that is, the domain

$$|\xi + \tau\rho^{-1}M_a| \equiv |\tau^{-1}z^0 + \rho^{-1}(z^1 + \cdots + z^n) + \tau\rho^{-1}C_a| < 1 \tag{2.11}$$

This inequality is satisfied, when C_a and ρ are given, for example by choosing τ such that

$$|\tau\rho^{-1}C_a| < \frac{1}{3} \tag{2.12}$$

and then restricting $\hat{\Omega}_1$ to the non-empty neighbourhood of $(0, \ldots, 0)$:

$$|z^0| < \frac{\tau}{3}, \quad |z^i| < \frac{\rho}{3n}. \tag{2.13}$$

We have proved, in particular under the condition (2.12), the existence of a real analytic solution of the problem in the polydisc $|x^i| < \frac{\rho}{3}$, for $|t| < \frac{\tau}{3}$. We see that the domain of existence does not depend on C_b nor $C_{\tilde{f}}$, and hence does not depend on C_f nor C_ϕ. □

Corollary 2.2 *(Global in space, local in time). Under the hypotheses of the theorem, and if the holomorphic extension of a is pointwise uniformly bounded*

[4] The various Cs denote the supremum of pointwise norms in Ω_1 of the holomorphic extensions of the quantity indicated in the index.

on $\hat{M} \times \hat{R}$, there is a neighbourhood $M \times \{|t| < \tau\}$ of M in $M \times R$ such that the system (2.1) has an analytic solution u taking on M the initial value ϕ.

Proof Glue together local solutions, using the uniqueness property and the paracompactness of M. □

Global in time existence can be proved, using the independence of the domain of local existence from C_f and C_ϕ

2.2 *Non-linear system*

A global in space, local in time theorem can also be proved for the analytic Cauchy problem for the quasilinear system, i.e. f linear[5] in the derivative $D_x u$,

$$\partial_t u = f(t, x, u, D_x u) \tag{2.14}$$

under appropriate holomorphy hypotheses of f, in t, x and also u. The solution can be constructed by iteration of solutions of linear systems.

3 Fuchs theorem

3.1 *Definitions*

We consider the partial differential first-order system on $V = M \times R$, $x \in M$, $t \in R$

$$t\partial_t u + A(t, x)u = tf(t, x, u, D_x u), \tag{3.1}$$

with A a linear operator analytic in x and continuous in t. We assume as in the previous section that f is linear in $D_x u$ and holomorphic in u for $|u| < c$. We still suppose that $\hat{M}$ extends to a complex analytic manifold $\hat{M}$ and that A extends to a linear operator and f to a mapping, both holomorphic in $z \in \hat{M}$. However, we suppose only that A and f are continuous in $t \in [0, T]$.

The Cauchy problem for the system (3.1) is singular for data at $t = 0$.

Definition 3.1 *The system (3.1) is called Fuchsian if there exist a number $\alpha < 1$ and a number $\Sigma > 0$ such that the linear operator $\sigma^{A(z)} := e^{A(z) log \sigma}$ satisfies for $t \in [0, T]$ the inequality*

$$Sup_{z \in \hat{M}} |\sigma^{A(t,z)}| \sigma^{\alpha} \leq \Sigma. \tag{3.2}$$

The following lemma results from elementary algebra.

Lemma 3.2 *The system (3.1) is Fuchsian if the linear operator A is uniformly bounded on $\hat{M} \times [0, T]$ with eigenvalues of real part greater than -1.*

[5] This property can be attained for more general systems by differentiating the equations and introducing derivatives as new unknowns.

Proof By definition one has

$$\sigma^A := e^{A\log\sigma} \equiv I + \sum_{k=1}^{\infty} \frac{1}{k!} A^k (\log\sigma)^k. \tag{3.3}$$

If $A(t,z)$ is uniformly bounded on $\hat{M} \times [0,T]$, it admits uniformly bounded eigenvalues $\lambda_1, \ldots, \lambda_N$, N the dimension on the vector space on which A acts. The representative of A is a matrix which admits a **Jordan decomposition**,

$$\begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & a_p \end{pmatrix}.$$

The matrix $\sigma^A := e^{A\log\sigma}$ admits then the block decomposition

$$e^{A\log\sigma} = \begin{pmatrix} e^{a_1\log\sigma} & 0 & 0 & 0 \\ 0 & e^{a_2\log\sigma} & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & e^{a_p\log\sigma} \end{pmatrix}. \tag{3.4}$$

A Jordan block a is an $m \times m$ triangular matrix associated with an eigenvalue of multiplicity m and reads, if for example $m = 3$

$$a = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}. \tag{3.5}$$

An exponential $e^{a\log\sigma}$ is

$$e^{a\log\sigma} \equiv I + a\log\sigma + \sum_{k=2}^{\infty} \frac{1}{k!} a^k (\log\sigma)^k, \tag{3.6}$$

with, by elementary calculus

$$a^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \\ & \lambda^k & \lambda^{k-1} \\ & & \lambda^k \end{pmatrix}$$

hence, again by elementary manipulations using the definitions

$$e^{a\log\sigma} = \begin{pmatrix} e^{\lambda\log\sigma} & e^{\lambda\log\sigma} - \log\sigma & 0 \\ & e^{\lambda\log\sigma} & e^{\lambda\log\sigma} - \log\sigma \\ & & e^{\lambda\log\sigma} \end{pmatrix}.$$

For $0 \le \sigma \le 1$, $\sigma^\alpha \log\sigma$ is bounded as soon as α is a positive number, while the function

$$e^{\lambda\log\sigma}\sigma^\alpha \equiv e^{(\lambda+\alpha)\log\sigma} \tag{3.7}$$

is bounded if the real part $\mathcal{R}e\lambda$ of λ is such that

$$\mathcal{R}e\lambda + \alpha > 0. \tag{3.8}$$

This inequality is satisfied for some $\alpha < 1$ if $\mathcal{R}e\lambda > -1$. The product $\sigma^\alpha e^{A\log\sigma}$ is bounded if the inequality is satisfied by all the eigenvalues of A. □

Remark 3.3 *The definition we give of a Fuchsian operator is less restrictive than the definition given by Kichenassamy–Rendall, which supposes that σ^A is bounded. It has the advantage to apply to operators with a multiple eigenvalue. The difficulty with the original definition was pointed out by Thibault Damour and Sophie de Beuyl in the case of nilpotent operators used in Hamiltonian systems*[6].

3.2 *Theorem*

We consider Equation (3.1), with

$$f(t, x, u, D_x u) \equiv f_0(t, x, u) + f_1(t, x, u) D_x u.$$

f_0 and f_1 are non-linear maps, which extend to holomorphic maps in $z \in \hat{M}$, and are holomorphic in u for $|u| < c$. They are continuous in $t \in [0, T]$.

Theorem 3.4 *If the system (3.1) is Fuchsian, then there exists a neighbourhood V of $M \times \{0\}$ in $M \times R$ such that this equation admits one and only one solution u, a tensor field on V analytic in $x \in M$, C^1 in t, and such that $u = 0$ for $t = 0$.*

Corollary 3.5 *If all the eigenvalues of A are such that*

$$\mathcal{R}e\lambda > -\mu, \tag{3.9}$$

then

1. *The existence theorem 3.4 holds for the equation*

$$t\partial_t u + Au = t^\mu f(t, x, u, \partial_x u), \quad \mu > 0, \tag{3.10}$$

2. *An analogous theorem holds, but with $u \to 0$ when $t \to +\infty$, for the equation*

$$\partial_t u - Au = e^{-\mu t} f(t, x, u, \partial_x u) \quad \mu > 0 \tag{3.11}$$

Proof of the corollary

1. The change of variable $\tau = t^\mu$ transforms Equation (3.10) into

$$\mu\tau \frac{d\tau}{dt} \partial_\tau u + Au = \tau f(\tau^{\mu^{-1}}, x, u, \partial_x u);$$

[6] Damour, T. and de Buyl, S. *Describing General Cosmological Singularities in Iwasawa Variables*, to be published.

that is,

$$\tau\partial_\tau u + \mu^{-1}Au = \tau\mu^{-1}f(\tau^{\mu^{-1}}, x, u, \partial_x u).$$

The theorem applies to this system, but its linear operator is $\mu^{-1}A$ whose eigenvalues are $\mu^{-1}\lambda$ if λ are the eigenvalues of A. The condition is therefore $\mu^{-1}\mathcal{R}e\lambda > -1$.

2. The change of variable $\tau = e^{-\mu t}$ transforms Equation (3.11) into

$$\tau\partial_\tau u + \mu^{-1}Au = -\mu^{-1}\tau f(\mu^{-1}\log\tau, x, u, \partial_x u). \tag{3.12}$$

□

Remark 3.6 *The corollary shows that the hypothesis on the boundedness of f in t when t tends to infinity can be replaced by a hypothesis that $e^{-\nu t}f$ is bounded for some $\nu > 0$ such that the eigenvalues of A satisfy the hypothesis, always implied by the original one (3.9).*

$$\mathcal{R}e\lambda > -\mu + \nu. \tag{3.13}$$

This remark permits the application of the corollary to functions f which have a polynomial growth as t tends to infinity. Corresponding statements hold in the case of the singularity for $t = 0$.

We now prove the theorem in several steps, first in a local chart, in which u is C^m valued.

3.3 *Equivalence with an integral equation*

We consider a chart of $\hat{M}$ whose image contains a ball of C^n, B_{s_0}, $|z| \leq s_0$. We keep the same notations for the fields and their representatives in the chart, a set of m complex functions taking real values for real z.

To solve Equation (3.1) extended to complex values

$$t\partial_t u + A(t,z)u = tf(t,z,u,\partial_z u) \equiv t\{f_0(t,x,u) + f_1(t,x,u)D_x u\}, \tag{3.14}$$

we set $u = Mv$, with M an $m \times m$ matrix satisfying the homogeneous equation

$$t\partial_t M + AM = 0.$$

We choose

$$M = t^{-A} \equiv e^{-A\log t}.$$

Equation (3.14) reads then

$$\partial_t v = t^A f(t, x, t^{-A}v, \partial_z(t^{-A}v)).$$

Hence Equation (3.14) together with the condition $u_{|t=0} = 0$ is equivalent to the integral equation

$$u(t,z) = t^{-A}\int_0^t \tau^A f(\tau, z, u(\tau,z), \partial_z u(\tau,z))d\tau;$$

i.e., setting $\tau = \sigma t$,

$$u(t,z) = t\int_0^1 \sigma^A f(\sigma t, z, u(\sigma t, z), \partial_z u(\sigma t, z))d\sigma. \tag{3.15}$$

We define[7] a Banach space $\mathcal{B}_a$ of C^m valued functions $v : (t,z) \mapsto v(t,z)$ holomorphic in z, real valued for z real, and continuous in t in the domain D of $B_{s_0} \times R$, depending on a real parameter $a > 0$,

$$D := \{t, z;\ |z| < s_0, \quad 0 \le t < a(s_0 - |z|)\}. \tag{3.16}$$

The norm in $\mathcal{B}_a$, called the a-norm, is given by

$$||v||_a := Sup_D \left\{ t^{-1}|v(t,.)|_s (s_0 - s) \left\{ 1 - \frac{t}{a(s_0 - s)} \right\}^{\frac{1}{2}} \right\}, \tag{3.17}$$

where we have set

$$|v(t,.)|_s := Sup_{|z|\le s}|v(t,z)|. \tag{3.18}$$

The definition of the a-norm implies for this s-norm

$$|v(t,.)|_s \le \frac{t}{s_0 - s} \left\{ 1 - \frac{t}{a(s_0 - s)} \right\}^{-\frac{1}{2}} ||v||_a. \tag{3.19}$$

3.4 *Equivalence with another mapping*

The a-norm cannot be used directly to solve Equation (3.5) by iteration, because the boundedness of $||v||_a$ does not imply the bounds of $|v|_s$ necessary for the holomorphy and estimate of the non-linear maps f_i. Kichenassamy and Rendall use the following artefact. They set

$$H(v)(t,z) := \int_0^1 \sigma^{-1}\sigma^A v(\sigma t, z)d\sigma \tag{3.20}$$

and remark that if v solves the equation

$$v(t,z) = G(v) := tf(t, H(v)(t,z), \partial_z H(v)(t,z)), \tag{3.21}$$

then $H(v)$ satisfies the integral equation (3.15). They continue the proof of the Fuchs theorem with the hypothesis that $|\sigma^A|$ is bounded. We have relaxed this hypothesis to (3.2) and we can show

Lemma 3.7 *The mapping $G : v \mapsto G(v)$ maps a ball $||v||_a \le R$ into itself if A satisfies the condition (3.2) and a is small enough.*

[7] Baouendi, S. and Goulaouic, C. (1977) *Commun. PDE*, **2**, 1151–62; Kichenassamy, S. and Rendall, A. (1998) *Class. Quant. Grav.*, **15**, 1339.

Proof For $t < a(s_0 - s)$, the definition of H and the inequality (3.17) between the a and s norms give

$$|H(v)(t,.)|_s \leq ||v||_a \Sigma \int_0^1 \sigma^{-\alpha}\sigma^{-1}\frac{\sigma t}{s_0 - s}\left\{1 - \frac{\sigma t}{a(s_0 - s)}\right\}^{-\frac{1}{2}} d\sigma. \tag{3.22}$$

One makes in the integral the change of variable from σ to ρ:

$$\sigma t = \rho a(s_0 - s), \quad \text{hence} \quad t d\sigma = a(s_0 - s)d\rho. \tag{3.23}$$

Then the above inequality reads

$$|H(v)(t,.)|_s \leq \Sigma ||v||_a a\varepsilon^{\alpha} \int_0^{\varepsilon} \rho^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho, \tag{3.24}$$

with

$$\varepsilon := \frac{t}{a(s_0 - s)}, \quad \text{i.e.} \quad \sigma \equiv \frac{\rho}{\varepsilon}, \qquad 0 \leq \varepsilon < 1. \tag{3.25}$$

To compute the integral in (3.24), we split it into two parts which we bound, using $0 \leq \frac{\varepsilon}{2} < \frac{1}{2}$,

$$\int_0^{\varepsilon} \rho^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho \leq \int_0^{\frac{\varepsilon}{2}} \rho^{-\alpha}\left\{1 - \frac{\varepsilon}{2}\right\}^{-\frac{1}{2}} d\rho + \int_{\frac{\varepsilon}{2}}^{\varepsilon} \left(\frac{\varepsilon}{2}\right)^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho$$

which, to simplify the writing of computations, we bound by

$$\int_0^{\varepsilon} \rho^{-\alpha}\left\{1 - \frac{\varepsilon}{2}\right\}^{-\frac{1}{2}} d\rho + \int_0^{\varepsilon} \left(\frac{\varepsilon}{2}\right)^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho. \tag{3.26}$$

We find then, using in the first integral $\frac{\varepsilon}{2} < \frac{1}{2}$ and $\alpha < 1$,

$$\int_0^{\varepsilon} \rho^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho \leq \frac{\sqrt{2}}{1 - \alpha}\varepsilon^{1-\alpha} + 2^{\alpha}\varepsilon^{-\alpha} 2\{1 - \sqrt{1 - \varepsilon}).$$

We denote by C any positive number independent of data and unknowns; the above inequality reads

$$\int_0^{\varepsilon} \rho^{-\alpha}\{1 - \rho\}^{-\frac{1}{2}} d\rho \leq C\varepsilon^{1-\alpha}. \tag{3.27}$$

The inequality (3.22) gives then

$$|H(v)(t,.)|_s \leq C\Sigma ||v||_a a\varepsilon, \quad \varepsilon = \frac{t}{a(s_0 - s)} < 1. \tag{3.28}$$

Choose the ball $||v||_a \leq R$ such that $|H(v)|_s < c$ when $s < s_0$ and $t < a(s_0 - s)$, i.e. $C\Sigma aR < c$. The mappings f_0 and f_1 defining the mapping f (see 3.14) are then holomorphic and uniformly bounded in absolute value by numbers M_0 and M_1. Therefore if $w := G(v)$

$$|w(t,.)|_s \leq t\{M_0 + M_1|\partial_z H(v)|\}. \tag{3.29}$$

By the definition of $H(v)$ we have

$$|\partial_z H(v)| \leq \int_0^1 \sigma^{-1}|\sigma^A|\ |\partial_z v(\sigma t, z)|_s d\sigma \tag{3.30}$$

To estimate $|\partial_z v(\sigma t, z)|$, use a Cauchy inequality which says that, for any s_σ such that $s < s_\sigma < s_0$, one has

$$|\partial_z v(\sigma t, .)|_s \leq C \frac{|v(\sigma t, .)|_{s_\sigma}}{s_\sigma - s}. \tag{3.31}$$

We deduce from (3.19) that, if t is such that

$$\sigma t < a(s_0 - s_\sigma), \tag{3.32}$$

we have

$$|v(\sigma t, .)|_{s_\sigma} \leq ||v||_a \frac{\sigma t}{s_0 - s_\sigma} \left\{1 - \frac{\sigma t}{a(s_0 - s_\sigma)}\right\}^{-\frac{1}{2}}, \tag{3.33}$$

We take as the number s_σ, for given s and σ,

$$s_\sigma := \frac{1}{2}\left(s + s_0 - \frac{\sigma t}{a}\right), \quad \text{i.e. } s_0 - s_\sigma = \frac{1}{2}\left(s_0 - s + \frac{\sigma t}{a}\right). \tag{3.34}$$

Then the inequality (3.32) for t becomes

$$\sigma t < \frac{1}{2}a(s_0 - s) + \frac{\sigma t}{2}, \quad \text{i.e. } \sigma t < a(s_0 - s),$$

and hence the inequality (3.19) on v holds for all $s < s_0$, $t < a(s_0 - s)$ and $0 \leq \sigma \leq 1$.

Using the expression of s_σ the inequality (3.31) becomes

$$|\partial_z v(\sigma t, .)|_s \leq ||v||_a \frac{\sigma t}{(s_0 - s)^2 - (\sigma t/a)^2} \left\{1 - \frac{\sigma t}{a(s_0 - s_\sigma)}\right\}^{-\frac{1}{2}}. \tag{3.35}$$

Elementary calculus gives

$$(s_0 - s)^2 - \left(\frac{\sigma t}{a}\right)^2 = (s_0 - s)^2(1 - \rho^2)$$

and, using (3.34),

$$1 - \frac{\sigma t}{a(s_0 - s_\sigma)} = 1 - \frac{2\sigma t}{a(s_0 - s + \frac{\sigma t}{a})} = \frac{1 - \rho}{1 + \rho} > \frac{1 - \rho}{2}. \tag{3.36}$$

Therefore

$$\left(1 - \frac{\sigma t}{a(s_0 - s_\sigma)}\right)^{-\frac{1}{2}} < \sqrt{2}(1 - \rho)^{-\frac{1}{2}}. \tag{3.37}$$

Elementary calculus gives, $0 \leq \rho \leq 1$,

$$\frac{1}{(1-\rho^2)\sqrt{1-\rho}} \leq \frac{1}{(1-\rho)^{\frac{3}{2}}}.$$

We then deduce from (3.30), using the hypothesis on σ^A

$$|\partial_z H(v)| \leq \Sigma ||v||_a \frac{a}{s_0 - s} \varepsilon^\alpha \int_0^\varepsilon \rho^{-\alpha} \frac{d\rho}{(1-\rho)^{\frac{3}{2}}} \tag{3.38}$$

We estimate the integral as we estimated (3.26). We find

$$\int_0^\varepsilon \rho^{-\alpha} \frac{d\rho}{(1-\rho)^{\frac{3}{2}}} \leq 2^{\frac{3}{2}} \int_0^\varepsilon \rho^{-\alpha} d\rho + \left(\frac{\varepsilon}{2}\right)^{-\alpha} \int_0^\varepsilon \frac{d\rho}{(1-\rho)^{\frac{3}{2}}} \tag{3.39}$$

with

$$\int_0^\varepsilon \frac{d\rho}{(1-\rho)^{\frac{3}{2}}} = 2\left\{\frac{1}{\sqrt{1-\varepsilon}} - 1\right\} = 2\frac{1-\sqrt{1-\varepsilon}}{\sqrt{1-\varepsilon}} \leq \frac{2\varepsilon}{\sqrt{1-\varepsilon}}. \tag{3.40}$$

Therefore

$$|\partial_z H(v)|_s < C\Sigma ||v||_a \frac{a}{s_0 - s} \frac{\varepsilon}{\sqrt{1-\varepsilon}}. \tag{3.41}$$

The definition of the a-norm gives

$$||t\partial_z H(v)||_a < Ca\Sigma ||v||_a. \tag{3.42}$$

Hence

$$||w||_a \leq a\{s_0 M_0 + C\Sigma M_1 ||v_a||\}. \tag{3.43}$$

We have proved that w belongs to the ball $||w||_a < R$ of the Banach space B_a, if a is small enough. □

3.5 *Convergence of iterations*

To prove the convergence of the iterations it is sufficient to prove that the mapping is contracting.

Lemma 3.8 *If a is small enough the mapping G is a contraction mapping on B_a.*

Proof The proof follows the same lines as in the previous section. By the hypothesis on the mappings f_i we have, if $V, v \in B_a$,

$$|f_0(H(V)) - f_0(H(v))| \leq M_0' |H(V - v)| \tag{3.44}$$

and

$$\begin{aligned} &|f_1(H(V)\partial_z H(V)) - f_1(H(v)\partial_z H(v)| \\ &\quad \leq M_1' |H(V - v)|\ |\partial_z H(V)| + M_1 |\partial_z H(V - v)|. \end{aligned} \tag{3.45}$$

with M_0' and M_1' the bounds of the derivatives with respect to u of f_0 and f_1 in the domain $|u| \leq c$.

We deduce from (3.44) and (3.45) the inequality, with $W = G(V)$ and $w = G(v)$,

$$||W - w||_a \leq M_0' I_a' + M_1 I I_a^{'} + M_1' I \tilde{I}_a^{'},$$

where we have set, the inequalities being consequence of results (3.28) and (3.42),

$$I_a' := ||tH(V - v)||_a \leq C s_0 \Sigma ||V - v||_a a.$$
$$I I_a' := ||t\partial_z H(V - v)||_a \leq C\Sigma a ||V - v||_a.$$

We consider the last term

$$I\tilde{I}_a = ||t|H(V - v)| \ |\partial_z H(V)| \ ||_a$$

By the definition of the a-norm

$$I\tilde{I}_a = \underset{s<s_0, t<a(s_0-s)}{Sup} \left\{ (s_0 - s)|H(V - v)(t, .)|_s \ |\partial_z H(V)(t, .)|_s \left(1 - \frac{t}{a(s_0 - s)}\right)^{\frac{1}{2}} \right\}.$$
$$\leq \left\{ \underset{s<s_0, t<a(s_0-s)}{Sup} |H(V - v)(t, .)|_s \right\} \left\{ \underset{s<s_0, t<a(s_0-s)}{Sup} (s_0 - s)|\partial_z H(V)|_s \left(1 - \frac{t}{a(s_0 - s)}\right)^{\frac{1}{2}} \right\}$$

Hence using the inequalities (3.28) and (3.42)

$$I\tilde{I}_a \leq C\Sigma^2 ||V||_a ||V - v||_a a$$

Assembling these results shows that the mapping G is contracting for small enough a. □

We have proven that the mapping G admits a fixed point $\bar{v} \in B_a$, if a is small enough, which can be constructed by iteration and is unique. The field $H(\bar{v})$ satisfies the integral equation (3.13) in the domain D, is real for real z, and vanishes for $t = 0$. It is a solution of the Fuchsian problem in the domain $|x| < s_0$, $0 < t < a(s_0 - |x|)$ of the considered chart of M.

3.6 *Global in space theorem*

To complete the proof of Theorem 3.2, one glues together local solutions by standard methods using the uniqueness of the solutions we have constructed in neighbourhoods $t < a(s_0 - s)$, as in the case of the Cauchy–Kovalevski theorem.

APPENDIX VI

CONFORMAL METHODS

1 Introduction

The causal structure of a spacetime depends on its Lorentzian metric only up to a conformal factor. Conformal properties play an important role in General Relativity, particularly in the study of asymptotic properties of the gravitational field.

2 Conformal metrics. Confomorphisms

We use indifferently Latin or Greek indices when the definitions and results do not depend on the signature of the pseudo-Riemannian metric, which we call simply "metric".

Definitions. Two metrics g and $\tilde{g}$ on a manifold V are called conformal if there is a positive scalar function Ω on V such that

$$\tilde{g} = \Omega^2 g, \quad \text{i.e.} \quad \tilde{g}_{\alpha\beta} = \Omega^2 g_{\alpha\beta}, \quad \tilde{g}^{\alpha\beta} = \Omega^{-2} g^{\alpha\beta}. \tag{2.1}$$

A diffeomorphism f of (V, g) onto $(\tilde{V}, \tilde{g})$ is called a confomorphism if the pull back of $\tilde{g}$ by f is conformal to g, i.e. there exists a positive scalar function Ω on V such that

$$f^* \tilde{g} = \Omega^2 g. \tag{2.2}$$

The following formulae are independent of the signature of the metric.

2.1 *Connections of conformal metrics*

The Christoffel symbols of two conformal metrics $\tilde{g} = \Omega^2 g$ are computed to be linked by the relations

$$\tilde{\Gamma}^{\alpha}_{\beta\gamma} = \Gamma^{\alpha}_{\beta\gamma} + S^{\alpha}_{\beta\gamma}, \tag{2.3}$$

where S is the tensor

$$S^{\alpha}_{\beta\gamma} := \Omega^{-1}(\delta^{\alpha}_{\gamma}\partial_{\beta}\Omega + \delta^{\alpha}_{\beta}\partial_{\gamma}\Omega - g_{\beta\gamma}\partial^{\alpha}\Omega). \tag{2.4}$$

To simplify computations, we set

$$\Omega := e^{\Phi}, \quad \text{i.e.} \quad .\tilde{g}_{\alpha\beta} = e^{2\Phi} g_{\alpha\beta} \tag{2.5}$$

The tensor S difference of the connections of the metrics $\tilde{g}_{\alpha\beta}$ and $g_{\alpha\beta}$ reads then

$$\tilde{\Gamma}^{\lambda}_{\beta\mu} - \Gamma^{\lambda}_{\beta\mu} := S^{\lambda}_{\beta\mu} \equiv \delta^{\lambda}_{\mu}\partial_{\beta}\Phi + \delta^{\lambda}_{\beta}\partial_{\mu}\Phi - g_{\beta\mu}\partial^{\lambda}\Phi. \tag{2.6}$$

2.2 *Riemann tensors of conformal metrics*

The expression of the Riemann tensor

$$R_{\alpha\beta}{}^{\lambda}{}_{\mu} = \partial_\alpha \Gamma^\lambda_{\beta\mu} - \partial_\beta \Gamma^\lambda_{\alpha\mu} + \Gamma^\lambda_{\alpha\rho}\Gamma^\rho_{\beta\mu} - \Gamma^\lambda_{\beta\rho}\Gamma^\rho_{\alpha\mu}, \tag{2.7}$$

this definition, and the fact that the coefficients of a connection (for instance Γ) can always be made zero at one point, lead easily to the formula

$$\tilde{R}_{\alpha\beta}{}^{\lambda}{}_{\mu} - R_{\alpha\beta}{}^{\lambda}{}_{\mu} \equiv \nabla_\alpha S^\lambda_{\beta\mu} - \nabla_\beta S^\lambda_{\alpha\mu} + S^\lambda_{\alpha\rho}S^\rho_{\beta\mu} - S^\lambda_{\beta\rho}S^\rho_{\alpha\mu}, . \tag{2.8}$$

One deduces then from (2.6) that

$$\tilde{R}_{\alpha\beta}{}^{\lambda}{}_{\mu} - R_{\alpha\beta}{}^{\lambda}{}_{\mu} \equiv f_{\alpha\beta}{}^{\lambda}{}_{\mu} - f_{\beta\alpha}{}^{\lambda}{}_{\mu} \tag{2.9}$$

with (indices raised with g)

$$f_{\alpha\beta}{}^{\lambda}{}_{\mu} := \delta^\lambda_\mu \nabla_\alpha \partial_\beta \Phi + \delta^\lambda_\beta \nabla_\alpha \partial_\mu \Phi - g_{\beta\mu}\nabla_\alpha \partial^\lambda \Phi \tag{2.10}$$

$$+ (\delta^\lambda_\alpha \partial_\rho \Phi + \delta^\lambda_\rho \partial_\alpha \Phi - g_{\alpha\rho}\partial^\lambda \Phi)(\delta^\rho_\beta \partial_\mu \Phi + \delta^\rho_\mu \partial_\beta \Phi - g_{\beta\mu}\partial^\rho \Phi). \tag{2.11}$$

We set as in CB-DM1 V

$$\Phi_{\alpha\beta} := \nabla_\alpha \partial_\beta \Phi - \partial_\alpha \Phi \partial_\beta \Phi \tag{2.12}$$

and we find, using $\Phi_{\alpha\beta} = \Phi_{\beta\alpha}$, that

$$\tilde{R}_{\alpha\beta}{}^{\lambda}{}_{\mu} - R_{\alpha\beta}{}^{\lambda}{}_{\mu} \equiv \delta^\lambda_\beta \Phi_{\alpha\mu} - \delta^\lambda_\alpha \Phi_{\beta\mu} + g_{\alpha\mu}\Phi_\beta{}^\lambda - g_{\beta\mu}\Phi_\alpha{}^\lambda + (\delta^\lambda_\beta g_{\alpha\mu} - \delta^\lambda_\alpha g_{\beta\mu})\partial_\rho \Phi \partial^\rho \Phi. \tag{2.13}$$

2.3 *Ricci tensors of conformal metrics*

Due to their appearance in different settings we give formulas corresponding to different choices of the form of the conformal factor.

2.3.1 *Case* $\tilde{g} = e^{2\Phi} g$

Contraction of the Riemann tensor (2.7) on α and λ indices gives for the Ricci tensor, $R_{\beta\mu}$, on a d-dimensional manifold[1],

$$\tilde{R}_{\beta\mu} - R_{\beta\mu} = -g_{\beta\mu}\nabla^\alpha \partial_\alpha \Phi - (d-2)\nabla_\mu \partial_\beta \Phi + (d-2)\{\partial_\mu \Phi \partial_\beta \Phi - g_{\beta\mu}\partial^\lambda \Phi \partial_\lambda \Phi\}. \tag{2.14}$$

Therefore, for the difference of the scalar curvatures the expression used in Chapter 7

$$e^{2\Phi}\tilde{R} - R = -2(d-1)\nabla^\alpha \partial_\alpha \Phi - (d-2)(d-1)\partial^\lambda \Phi \partial_\lambda \Phi. \tag{2.15}$$

[1] Remark that $\nabla^\alpha \partial_\alpha$is the Laplacian Δ_g if g is Riemannian, the wave operator $\Box_g$ if g is Lorentzian.

2.3.2 *Case* $\tilde{g} = \Omega^2 g$

If $\Phi = \log \Omega$, then

$$\partial_\alpha \Phi = \Omega^{-1} \partial_\alpha \Omega, \quad \nabla_\beta \partial_\alpha \Phi = \Omega^{-1}(\nabla_\beta \partial_\alpha \Omega - \Omega^{-1} \partial_\beta \Omega \partial_\alpha \Omega), \tag{2.16}$$

the formulae (2.12) and (2.13) are found to become

$$\begin{aligned} \tilde{R}_{\alpha\beta} - R_{\alpha\beta} = &-\Omega^{-1}\{g_{\alpha\beta}\Delta_g \Omega + (d-2)\nabla_\alpha \partial_\beta \Omega\} + \\ &+ \Omega^{-2}\{2(d-2)\partial_\alpha \Omega \partial_\beta \Omega - (d-3)\ g_{\alpha\beta}\partial^\lambda \Omega \partial_\lambda \Omega\}, \end{aligned} \tag{2.17}$$

and, using the elementary identity

$$2(d-2) - d(d-3) \equiv -(d-4)(d-1), \tag{2.18}$$

$$\Omega^2 \tilde{R} - R = -2(d-1)\Omega^{-1}\Delta_g \Omega - (d-4)(d-1)\Omega^{-2}\partial^\lambda \Omega \partial_\lambda \Omega. \tag{2.19}$$

Remark that the quadratic term in $\partial\Omega$ vanishes in dimension $d = 4$.

2.3.3 *Case* $d = 3$, $\tilde{g} = \phi^4 g$

This case is particularly important for everyday physics, and enjoys many special geometric properties. In this case (2.17) reduces to

$$\tilde{R}_{\alpha\beta} - R_{\alpha\beta} = -\Omega^{-1}\{g_{\alpha\beta}\nabla^\lambda \partial_\lambda \Omega + \nabla_\alpha \partial_\beta \Omega\} + 2\Omega^{-2}\partial_\alpha \Omega \partial_\beta \Omega. \tag{2.20}$$

We have seen in Chapter 7 that the choice $\tilde{g} = \phi^4 g$ leads to a relation between the scalar curvatures of g and $\tilde{g}$ free of quadratic terms in the derivative of ϕ. We compute the relation between the Ricci tensors, setting $\Omega = \phi^2$ in (2.20) i.e.

$$\partial_\alpha \Omega = 2\phi \partial_\alpha \phi, \quad \nabla_\beta \partial_\alpha \Omega = 2\phi \nabla_\beta \partial_\alpha \phi + 2\partial_\beta \phi \partial_\alpha \phi. \tag{2.21}$$

We find, in the case $d = 3$

$$\tilde{R}_{\alpha\beta} - R_{\alpha\beta} = -\phi^{-2}\{g_{\alpha\beta}(2\phi\Delta_g \phi + 2\partial_\lambda \phi \partial^\lambda \phi) + 2\phi\nabla_\beta \partial_\alpha \phi - 6\partial_\alpha \phi \partial_\beta \phi\} \tag{2.22}$$

and the relation used in Chapter 7 with a Riemannian metric g,

$$\phi^4 \tilde{R} - R = -8\phi^{-1}\Delta_g \phi. \tag{2.23}$$

3 The Weyl tensor

The formulae obtained for the Riemann, Ricci, and scalar curvature of two conformal metrics show that the following tensor W, called the **Weyl tensor** is the same for conformal metrics

$$\begin{aligned} W_{ij,kl} = R_{ij,kl} &- \frac{1}{d-2}(g_{jl}R_{ki} - g_{il}R_{kj} + g_{jk}R_{li} - g_{ik}R_{jl}) \\ &- \frac{1}{(d-1)(d-2)}(g_{jk}g_{li} - g_{ik}g_{jl})R. \end{aligned} \tag{3.1}$$

It is considered that the Weyl tensor embodies in some sense the non-Newtonian properties[2] of the gravitational field, in particular its radiation

[2] This point of view is supported by the fact that the equations for massless fields, at least in four spacetime dimensions, are conformally invariant.

properties, the Weyl tensor is used to classify such properties. It is used also in global existence proofs, in particular in the proof of the non-linear stability of Minkowski spacetime (see Chapters 15 and 16).

The Weyl tensor has the same symmetries as the Riemann tensor. It is equal to the Riemann tensor in a Ricci flat space ($R_{ij} \equiv 0$). It has a zero trace.

Since W is obviously zero for a flat metric, it is also zero if the metric is conformal to a flat metric. It can be proved that if $d > 3$ the identical vanishing of the Weyl tensor implies that the metric is locally conformally flat.

The definition (3.1) shows that the Weyl tensor of a three-dimensional[3] Riemannian manifold (M, g) is identically zero.

It can be proved that a 3-manifold is locally conformally flat if and only if its **Coton**[4] **tensor** vanishes. This tensor is a 3-tensor with components given in terms of the Ricci tensor and the scalar curvature of g by

$$R_{kij:} := \nabla_i R_{jk} - \nabla_j R_{ik} + \frac{1}{4}(g_{ki}\partial_j R - g_{kj}\partial_i R). \tag{3.2}$$

4 Conformal transformations of field equations

4.1 *Maxwell and Yang–Mills equations*

Theorem 4.1 *The Maxwell equations in vacuum on a four-dimensional Lorentzian manifold (V, g) are invariant by confomorphisms.*

Proof The Maxwell equations for the electromagnetic 2-form F are, in the absence of charged sources, the metric independent equation

$$dF = 0 \tag{4.1}$$

and, with δ the codifferential operator,

$$\delta F = 0, \quad \text{i.e.} \quad \nabla_\alpha F^{\alpha\beta} = 0. \tag{4.2}$$

These equations are invariant by diffeomorphisms. On the other hand, straightforward computation shows that if

$$g_{\alpha\beta} = \Omega^{-2}\tilde{g}_{\alpha\beta}, \quad \text{hence} \quad g^{\alpha\beta} = \Omega^2\tilde{g}^{\alpha\beta} \tag{4.3}$$

and

$$F_{\alpha\beta} = \tilde{F}_{\alpha\beta} \quad \text{hence} \quad F^{\alpha\beta} = \Omega^4\tilde{F}^{\alpha\beta}, \tag{4.4}$$

then, in the case of a spacetime of dimension 4, the relation between connections implies

$$\nabla_\alpha F^{\alpha\beta} \equiv \Omega^4\tilde{\nabla}_\alpha\tilde{F}^{\alpha\beta}. \tag{4.5}$$

The theorem follows. □

[3] Recall that two-dimensional manifolds are all locally conformally flat.

[4] Coton (1926); see e.g. Eisenhart (1949) *Riemannian Geometry*, Princeton University Press.

Corollary 4.2 *The Yang–Mills equations on a four-dimensional manifold are conformally invariant.*

The proof, similar to the previous one, is left to the reader. This conformal invariance was used by Choquet-Bruhat and Christodoulou to prove global existence for small data (see Chapter 15).

5 Invariance of wave equations

A kind of invariance of the wave equation by confomorphisms holds for arbitrary spacetime dimension, but is more subtle.

Theorem 5.1 *Let g and $\tilde{g}$ be two conformal metrics on a d-dimensional manifold V, $\tilde{g} = \Omega^{-2} g$. The wave equations in these metrics satisfy the identity, invariant by diffeomorphisms:*

$$\nabla^\alpha \partial_\alpha u - \frac{d-2}{4(d-1)} R u \equiv \Omega^{-(2+d)/2} \left(\tilde{\nabla}^\alpha \partial_\alpha \tilde{u} - \frac{d-2}{4(d-1)} \tilde{R} \tilde{u} \right) \tag{5.1}$$

where R and $\tilde{R}$ are the scalar curvatures of g and $\tilde{g}$ while $\tilde{u}$ is such that

$$\tilde{u} = \Omega^{(d-2)/2} u \tag{5.2}$$

Proof Straightforward computation[5] □

Remark 5.2 *The identity (5.1) is a rewriting of an identity written in Chapter 7, with $\Omega = \varphi^{\frac{2}{d-2}}$.*

6 Penrose transform

Penrose has defined a confomorphism from the four-dimensional Minkowski spacetime M^4 onto an open bounded subset of the Einstein cylinder $\Sigma^4 := S^3 \times R$.

A number of global existence results[6] have been obtained for conformally invariant field equations on Minkowski spacetime by using the Penrose conformal compactification.

The **Penrose confomorphism** extends as follows to arbitrary spacetime dimensions.

Let $r, \theta^1, \ldots, \theta^{n-1}$ be polar (pseudo) coordinates on R^n, i.e. $r \geq 0$, $\theta^1, \ldots, \theta^{n-1}$ angular coordinates on S^{n-1}. In such coordinates the metric of the Minkowski spacetime M^{n+1} reads, with $t \in R$

$$\eta \equiv -dt^2 + dr^2 + r^2 (dS^{n-1})^2 \tag{6.1}$$

$$(dS^{n-1})^2 = (d\theta^1)^2 + \sin^2 \theta^1 [(d\theta^2)^2 + \cdots + (\sin^2 \theta^2 \sin^2 \theta^3 \ldots \sin^2 \theta^{n-1})(d\theta^{n-1})^2] \tag{6.2}$$

[5] See CB - DM I V problem 4.
[6] See Chapter 15.

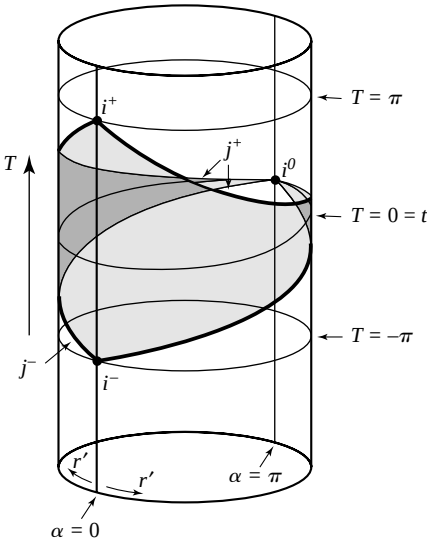


Figure F.1 Image of Minkowski spacetime in Einstein cylinder.

Let $\alpha, \Theta^1, \ldots, \Theta^{n-1}$ be the usual angular (pseudo) coordinates on S^n, i.e. $0 \leq \alpha \leq \pi$ and Θ^i angular coordinates on S^{n-1}. The subsets $\alpha =$ constant are spheres S^{n-1}, except $\alpha = 0$ and $\alpha = \pi$ which reduce to a point. The metric of $\Sigma^{n+1} := S^n \times R$ is, with $T \in R$

$$\gamma \equiv -dT^2 + d\alpha^2 + \sin^2 \alpha (dS^{n-1})^2 \tag{6.3}$$

We consider the mapping $\mathbf{P}$ from M^{n+1} into Σ^{n+1} given by

$$\Theta^i = \theta^i, \quad i = 1, \ldots, n-1, \tag{6.4}$$

$$T = tg^{-1}u + tg^{-1}v, \quad u =: t + r, \quad v =: t - r \tag{6.5}$$

$$\alpha = tg^{-1}u - tg^{-1}v \tag{6.6}$$

The range of this mapping is the open bounded set $\mathcal{V} \subset \Sigma^{n+1}$ defined by

$$0 \leq \alpha < \pi, \quad (\theta^i) \in S^{n-1}, \quad \alpha - \pi < T < \pi - \alpha. \tag{6.7}$$

The submanifold $t = 0$ of M^{n+1} is mapped onto the sphere $S^n \times \{\, T = 0\}$ minus the point $\alpha = \pi$. This point, which can be considered as the image under the Penrose map of initial space infinity ($t = 0, r = +\infty$), is traditionally named I_0.

The boundary of $\mathcal{V}$ in Σ^{n+1} consists of the point I_0, two points I_+, and I_- called respectively future and past time infinity, with coordinates respectively $\alpha = \pi, T = \pi$ and $\alpha = \pi, T = -\pi$, and two portions of null conoids hypersurfaces $\mathcal{I}_+$ and $\mathcal{I}_-$, called respectively future and past null infinity, with equations $T = \pi - \alpha > 0$ and $T = \alpha - \pi < 0$ with vertex respectively I_+ and I_-.

Theorem 6.1 *The Penrose mapping* $\boldsymbol{P}$ *is a confomorphism from* M^{n+1} *onto the bounded open set* $\mathcal{V} \subset \Sigma^{n+1}$*; it holds that*

$$\mathbf{P}^*\eta = \Omega^{-2}\gamma, \quad \textit{with} \quad \Omega = \cos T + \cos\alpha. \tag{6.8}$$

Proof The mapping $\mathbf{P}$ is a diffeomorphism with inverse

$$t = \frac{1}{2}\left\{tg\frac{T+\alpha}{2} + tg\frac{T-\alpha}{2}\right\}, \quad r = \frac{1}{2}\left\{tg\frac{T+\alpha}{2} - tg\frac{T-\alpha}{2}\right\}; \tag{6.9}$$

equivalently, using trigonometric identities

$$r = (\cos T + \cos\alpha)^{-1}\sin\alpha, \quad t = (\cos T + \cos\alpha)^{-1}\sin T. \tag{6.10}$$

Straightforward calculus gives then

$$dt = \frac{(1+\cos T\cos\alpha)dT + \sin T\sin\alpha d\alpha}{(\cos T + \cos\alpha)^2} \tag{6.11}$$

$$dr = \frac{(1+\cos T\cos\alpha)d\alpha + \sin T\sin\alpha dT}{(\cos T + \cos\alpha)^2} \tag{6.12}$$

From these formulae one deduces the equality

$$-dt^2 + dr^2 = (\cos T + \cos\alpha)^{-2}(-dT^2 + d\alpha^2). \tag{6.13}$$

The theorem is a consequence of this equality. □

Remark 6.2 *The function* $\Omega \equiv \cos T + \cos\alpha$ *is positive on* $\mathcal{V}$*. it vanishes on the boundary* $\partial\mathcal{V}$ *of* $\mathcal{V}$ *in* Σ^{n+1}*.*

Remark 6.3 *The relation (6.8) reads, equivalently:*

$$\eta = [\Omega \circ \mathbf{P}]^{-2}\Phi'\gamma), \tag{6.14}$$

with in polar coordinates on R^n

$$[\Omega \circ \mathbf{P}]^2 = \frac{4}{\{1+(t+r)^2\}\{1+(t-r)^2\}} \tag{6.15}$$

The following proposition, a consequence of the definition of the Penrose mapping, is particularly important, because it is taken in a generalized context as characteristic of strong asymptotically Minkowskian spacetimes through some conformal mapping.

Proposition 6.4 *The images in* Σ^{n+1} *by the Penrose mapping of submanifolds* $t =$*constant*[7] *of* M^{n+1} *are spacelike and turn null at the boundary of* $\mathcal{V}$*, except for* $t = 0$*. The reciprocal images of the submanifolds* $T =$ *constant* $\cap\,\mathcal{V}$ *are spacelike hyperboloids of Minkowski space which turn null at infinity. The images of Minkowskian future [past] timelike lines all converge in* Σ^{n+1} *to* I_+ $[I_-]$*. Images of null lines have end points either on* $\mathcal{I}_+$ *or on* $\mathcal{I}_-$*.*

[7] Or strictly spacelike submanifolds.

7 Einstein spaces with cosmological constant

The model vacuum Einsteinian spaces with cosmological constant, De Sitter and anti-De Sitter spaces, have very different properties from the Minkowski spacetime, and are very different from each other.

7.1 *Conformal transformation of De Sitter spacetime*

The De Sitter spacetime, $(S^3 \times R, g_{\text{de Sitter}})$ is a solution of the vacuum Einstein equations with positive cosmological constant

$$g_{\text{de Sitter}} := -dt^2 + k^{-2} cosh^2(kt)\{d\alpha^2 + \sin^2\alpha(d\theta^2 + \sin^2\theta d\phi^2)\}. \tag{7.1}$$

It is mapped onto a bounded strip of the Einstein cylinder Σ^4 by the confomorphism preserving the timelines

$$t' = 2\arctan\exp(k^{-1}t) - \frac{1}{2}\pi. \tag{7.2}$$

7.2 *Conformal transformation of anti-De Sitter spacetime*

The anti-de Sitter spacetime is a solution of the vacuum Einstein equations with negative cosmological constant

$$g_{antiDesit} \equiv -dt^2(\cosh\chi)^2 + d\chi^2 + \sinh^2\chi(\sin^2\theta d\phi^2 + d\theta^2) \equiv (\cosh\chi)^2 g_{\Sigma^4}. \tag{7.3}$$

It is mapped onto a timelike half of the Einstein cylinder, $0 \leq \alpha < \frac{\pi}{2}$, by the confomorphism

$$\alpha = 2\arctan\ \exp\chi - \frac{1}{2}\pi.$$

8 Asymptotically simple spacetimes

The conformal mapping of Minkowski spacetime onto a bounded open subset of the Einstein cylinder has led Penrose and followers to define asymptotically Minkowskian spacetimes through the use of more general conformal mappings, supposed to exist. Such definitions are used extensively in the current literature to characterize the asymptotic behaviour of gravitational fields.

8.1 *Conformal compactifications*

There exist in the literature several definitions of the conformal compactification of a spacetime through its image under a confomorphism. The most general definition is as follows.

Definition 8.1 *A conformal compactification of a spacetime (V, g) is a confomorphism f of (V, g) onto $(\tilde{U}, \tilde{g})$ with $\tilde{U}$ a relatively compact, connected, open set with boundary $\partial\tilde{U}$ in a spacetime $(\tilde{V}, \tilde{g})$,*

$$\tilde{g} = \Omega^2 f^* g \quad \text{in } \tilde{U}, \tag{8.1}$$

with a conformal factor Ω *such that*

$$\Omega > 0 \quad in \quad \tilde{U}, \tag{8.2}$$

$$\Omega = 0 \quad and \quad \partial\Omega \neq 0 \quad on \quad \partial\tilde{U}. \tag{8.3}$$

This definition does not state the differentiability properties of Ω, nor the properties of the boundary $\partial\tilde{U}$.

The following definition leaves open the true problem by including the undefined word "regular".

Definition 8.2 *A spacetime is called asymptotically simple if it admits a regular conformal compactification.*

The condition that Ω is a smooth function on the whole of $\tilde{V}$ essentially rules out asymptotically Euclidean spaces[8] with positive ADM mass.

The definition of asymptotically simple spacetimes has been weakened in various ways to accommodate realistic situations[9]. The first hypothesis to relax is the smoothness of the conformal factor at space infinity, if this is represented by a point i_0. A current definition of (future) asymptotic simplicity is that every inextendible future null geodesic has an end point of a part of $\partial\tilde{U}$ called $\mathcal{I}_+$. Spacetimes asymptotically simple in this sense have been constructed by Friedrich, and by Chruściel and Delay[10]. The smoothness of $\mathcal{I}_+$, assumed to exist, is related with the peeling property (asymptotic expansion of the Riemann tensor at space infinity), which is not yet proved in generality. The conformal compactification is an interesting geometrical construction, but made in non-physical spacetime, and leaves open many problems on the asymptotic behaviour of the physical metric[11].

The conformal compactification method to transform global problems on Minkowski spacetime into local existence problems on the Einstein cylinder for asymptotically Minkowskian solutions of conformally invariant fields equations has been generalized, introducing the conformal factor as another unknown, and used by H. Friedrich to prove, in a long series of papers, global existence and properties of solutions of the Einstein equations satisfying hyperboloidal initial conditions (see Chapter 15).

We send the reader interested in geometric conformal structure at space infinity to the article of Friedrich[12] and references therein.

[8] But it can be used for spacetimes with a positive cosmological constant, see e.g. Galloway, G. (2002) gr-qc 0208092v1.

[9] See Claudel, C. (2000) gr-qc 0005031v2.

[10] Chruściel, P. and Delay, E. (2002) gr-qc 0208092v1.

[11] On gravitational radiation see Damour, T. (1986) *Proc. Fourth Marcel Grossmann Meeting* (ed. R. Ruffini), Elsevier, pp. 365–92.

[12] Friedrich, H. (2004) in Friedrich, H. and Chruściel, P. (eds) (2004), Birkhäuser.

8.2 *Black holes*

Asymptotically Minkowskian black holes are often defined in the literature by supposing that the spacetime admits a compactification and the boundary $\partial\tilde{U}$ contains as in the original Penrose diagram subsets $\mathcal{I}_+$ and $\mathcal{I}_-$, which represent points of future and past null infinity but $\mathcal{I}_+$ does not contain an end point of every inextendible null geodesic. Spacetimes admitting such a compactification are often taken as a definition of spacetimes containing a black hole. The black hole region is defined as follows.

Definition 8.3 *The image of the black hole region $\mathcal{B}$ of a spacetime admitting a Penrose compactification $(U, \tilde{g})$ is the complementary in $\tilde{U}$ of the past of $\mathcal{I}_+$.*

*The **event horizon** in the original spacetime is the reciprocal image of the boundary $\partial\mathcal{B}$ of $\mathcal{B}$ in U.*

Many papers on black holes use these definitions, with various additional hypotheses. In particular it is generally supposed that the event horizon is a null submanifold[13].

[13] Though it is not necessarily the case under the previous assumptions; see Chruściel, P. (2006) gr-qc 0606101.

APPENDIX VII

KALUZA–KLEIN THEORIES

1 Introduction

Already in the 1920s Kaluza and Klein, in search of unification of gravitation with electromagnetism, considered the Einstein equations on a five-dimensional Lorentzian manifold with a 1-parameter isometry group of spatial isometries. They supposed that the Killing vector ξ has a constant length and neglected the equation $R(\xi,\xi)=0$. They showed that the 14 remaining Einstein equations *in vacuo* for $(V_5,\hat{g})$ split into Maxwell equations for a 2-form F, and Einstein equations for a spacetime (V_4,g) with source the corresponding Maxwell tensor. Later Jordan relaxed the hypothesis of the constancy of the length of ξ, interpreting the 15th potential as a varying gravitational constant. Shortly afterwards, Thiry, a student of Lichnerowicz, gave a geometric formulation of the equations. They interpreted the 15th potential as the induction of vacuum.

In 1968, Kerner indicated how the same technique applied to a d-dimensional spacetime with $d-4$-dimensional isometry group can lead to the coupled Einstein–Yang–Mills system.

2 Isometries

Recall that an **isometry** of a pseudo-Riemannian manifold $(\hat{V},\hat{g})$ is a diffeomorphism of $\hat{V}$ which leaves $\hat{g}$ invariant; see for instance CB-DM1 III D. Isometries of $(\hat{V},\hat{g})$ constitute a Lie group G_r of dimension r with

$$r \leq \frac{1}{2}d(d+1) \quad \text{if} \quad \hat{V} \text{ has dimension } d \tag{2.1}$$

A one-parameter group of diffeomorphisms of V generated by a vector field X is a group of isometries (subgroup of G_r) if and only if the Lie derivative of g with respect to X vanishes:

$$\mathcal{L}_X g = 0, \quad \text{i.e.} \quad \nabla_\alpha X_\beta + \nabla_\beta X_\alpha = 0. \tag{2.2}$$

The **orbit** $\mathcal{O}_x$ of a point x under the group G_r is the set of all points of V into which x is mapped by all elements of G_r. If the group acts **simply transitively** on $\mathcal{O}_x$, this orbit is a submanifold of V of dimension r.

3 Kaluza–Klein metrics

Let $(\hat{V}, \hat{g})$ be a pseudo-Riemannian manifold of dimension $\hat{d}$ and let G be a Lie group of dimension d of isometries of $(\hat{V}, \hat{g})$, i.e. such that the right action of G on $\hat{V}$ leaves $\hat{g}$ invariant. We suppose this action of G on $\hat{V}$ to be such that it endows $\hat{V}$ with the structure of a G-principal fibre bundle whose fibres are the orbits of G. These orbits are then all diffeomeorphic to G and G acts simply transitively on them. The base is a $n+1 = \hat{d} - d$-dimensional manifold V. We denote by π the bundle projection $\pi : \hat{V} \to V$. It is proved in CB-DM2 V 13 that, under these hypotheses, the datum of the metric $\hat{g}$ on $\hat{V}$ is equivalent to the data of a $G-$invariant metric Φ on each orbit of G (if these orbits are not null for $\hat{g}$), a G-connection on $\hat{V}$ whose horizontal spaces are the vector spaces orthogonal to the orbits (hence invariant under the action of G), and a pseudo-Riemannian metric g on the base V.

3.1 *Metric in adapted frame*

We consider a local trivialization ϕ of $\hat{V}$ over the domain U of a chart in V; that is, a mapping ϕ,

$$\phi : \pi^{-1}(U) \to U \times G. \tag{3.1}$$

We choose local coordinates $\hat{x}^M$ in $\hat{V}$ adapted to the trivialization i.e. such that $\hat{x}^\alpha = x^\alpha$, local coordinates in U, for $\alpha = 0, 1 \ldots, n$, and $\hat{x}^m = y^m$, $m = n+1, \ldots, \hat{d}$, local coordinates in G. We take a moving frame $(\hat{e}_\alpha, \hat{e}_m)$ such that the $\hat{e}_m$ are Killing vector fields of the right action of G on $\hat{V}$; they are mapped by ϕ onto a basis e_m of the right Killing vector fields on G. We choose the $\hat{e}_\alpha$ orthogonal[1] to the $\hat{e}_m$ and mapped by ϕ onto a basis e_α of the tangent space to $U \subset V$. Then

$$\phi' \hat{e}_m = e_m, \quad \phi' \hat{e}_\alpha = e_\alpha - A^m_\alpha e_m. \tag{3.2}$$

We denote by $\hat{\theta}^A$ the dual coframe of $\hat{e}_A$ on $\hat{V}$, by θ^α and θ^m the dual coframes respectively of e_α on U and e_m on G. We find that

$$\hat{\theta}^\alpha = \theta^\alpha \quad \text{and} \quad \hat{\theta}^m \equiv \theta^m + A^m_\alpha \theta^\alpha, \tag{3.3}$$

One shows that $A := A^m_\alpha \theta^\alpha \otimes e_m$ is the representant on V, in the trivialization ϕ, of a G-connection 1-form which takes its values in the Lie algebra $\mathcal{G}$ of G.

With this choice of frames the metric $\hat{g}$ of $\hat{V}$ can be written, in $\pi^{-1}(U)$,

$$\hat{g} \equiv \hat{g}_{AB} \hat{\theta}^A \hat{\theta}^B \equiv g_{\alpha\beta} \theta^\alpha \theta^\beta + \Phi_{mn} (\theta^m + A^m_\alpha \theta^\alpha)(\theta^n + A^n_\beta \theta^\beta) \tag{3.4}$$

where $g \equiv g_{\alpha\beta} \theta^\alpha \theta^\beta$ is a metric on V and $\Phi := (\Phi_{mn})$ a set of functions on V, sometimes called a scalar multiplet. Since the Lie derivatives of the 1-forms θ^α and θ^m in the direction of a Killing vector $\hat{e}_m$ vanish, the metric $\hat{g}$ is invariant

[1] The vectors $\hat{e}_\alpha$ are a basis of the horizontal space of the G-connection.

under the right action of G on $\hat{V}$ if and only if the coefficients $g_{\alpha\beta}$, Φ_{mn}, A^m_α depend only on the local coordinates x^α on V.

The metric $g \equiv g_{\alpha\beta}\theta^\alpha\theta^\beta$ on V, the connection A and the set of functions $\Phi := (\Phi_{mn})$ are geometric elements on the basis[2] V. Under a change of local trivialization of the bundle $\hat{V} \to V$ described by a transition mapping[3] $U \to G$ by $x \mapsto h(x)$ it holds that:

The metric g is invariant.

The connection A transforms as a representant of a G-connection,

$$A \to h^{-1}Ah + h^{-1}dh. \tag{3.5}$$

The scalar multiplet transforms as

$$\Phi \to h \otimes h\ \Phi. \tag{3.6}$$

3.2 *Structure coefficients*

The **structure coefficients** $\hat{c}^A_{BC}$ of a moving frame $\hat{\partial}_A$ with dual frame $\hat{\theta}^A$ are defined by the relations

$$[\hat{\partial}_B, \hat{\partial}_C] = \frac{1}{2}\hat{C}^A_{BC}\hat{\partial}_A, \quad d\hat{\theta}^A = -\frac{1}{2}\hat{C}^A_{BC}\hat{\theta}^B \wedge \hat{\theta}^C. \tag{3.7}$$

To simplify the computations we choose for θ^α a natural frame on V, $\theta^\alpha = dx^\alpha$. Then

$$d\hat{\theta}^\alpha \equiv d\theta^\alpha = 0, \quad \text{hence} \quad \hat{c}^\alpha_{BC} = 0 \tag{3.8}$$

Since the θ^m are dual to Killing vectors on G it holds that

$$d\theta^m = -\frac{1}{2}C^m_{pq}\theta^p \wedge \theta^q \tag{3.9}$$

where the C^m_{pq} are structure constants of the group G. We deduce from the expression of θ^m

$$d\hat{\theta}^m = \frac{1}{2}(-C^m_{pq}\theta^p \wedge \theta^q + (\partial_\alpha A^m_\beta - \partial_\beta A^m_\alpha)dx^\alpha \wedge dx^\beta). \tag{3.10}$$

The relation $\theta^m = \hat{\theta}^m - A^m_\alpha dx^\alpha$ and the antisymmetry of C^m_{pq} imply that

$$d\hat{\theta}^m = \frac{1}{2}(-C^m_{pq}\hat{\theta}^p \wedge \hat{\theta}^q + 2C^m_{pq}A^p_\alpha dx^\alpha \wedge \hat{\theta}^q + F^m_{\alpha\beta}dx^\alpha \wedge dx^\beta) \tag{3.11}$$

with

$$F^m_{\alpha\beta} \equiv \partial_\alpha A^m_\beta - \partial_\beta A^m_\alpha + C^m_{pq}A^p_\alpha A^q_\beta. \tag{3.12}$$

[2] The basis V is the quotient of $\hat{V}$ by the isometry group. It can be identified with a submanifold of V only if the family of horizontal spaces is completely integrable, i.e. tangent to an $n+1$-dimensional submanifold of $\hat{V}$.

[3] For notations and proofs see CB-DM1 IIIB 2.

We see from the definitions of the curvature of a connection on a principal fibre bundle (see CB-DM1 Vbis 4) that the $F^m_{\alpha\beta}$ are the components of the representative in the chart of V of the **curvature** of the **connection** A, since the curvature of A is the $\mathcal{G}$ valued antisymmetric tensor of type Ad[4]

$$F := \frac{1}{2}F^m_{\alpha\beta}dx^\alpha \wedge dx^\beta \otimes e_m \equiv dA + [A, A]. \tag{3.13}$$

We read on (3.11) that

$$\hat{C}^m_{np} = C^m_{np}, \quad \hat{C}^m_{\alpha\beta} = -F^m_{\alpha\beta}, \quad \hat{C}^m_{\alpha q} = -C^m_{pq}A^p_\alpha \tag{3.14}$$

3.3 *Kaluza–Klein connection*

Recall that (Chapter 6) the **connection coefficients** $\hat{\omega}^A_{BC}$ of a metric $\hat{g}$ are given in terms of the metric components $\hat{g}_{AB}$ and the structure coefficients by

$$\hat{\omega}^A_{BC} \equiv \hat{\Gamma}^A_{BC} + \tilde{\omega}^A_{BC},$$

where $\hat{\Gamma}^A_{BC}$ are the Christoffel symbols constructed with the coefficients $\hat{g}_{AB}$ and where

$$\tilde{\omega}^A_{BC} \equiv \frac{1}{2}(\hat{C}^A_{BC} - \hat{g}_{CL}\hat{g}^{MA}\hat{C}^L_{BM} - \hat{g}_{BL}\hat{g}^{MA}\hat{C}^L_{CM}) \tag{3.15}$$

We find easily that

$$\hat{\Gamma}^\alpha_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma}, \quad \tilde{\omega}^\alpha_{\beta\gamma} = 0, \tag{3.16}$$

where $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel symbols of the metric g on V,

$$\hat{\Gamma}^m_{pq} = \hat{\Gamma}^m_{\alpha\beta} = \hat{\Gamma}^\alpha_{\beta q} = 0 \tag{3.17}$$

and, with Φ^{mn} the inverse matrix of Φ_{mn},

$$\hat{\Gamma}^m_{\alpha q} = \frac{1}{2}\Phi^{mn}\partial_\alpha\Phi_{nq}, \quad \hat{\Gamma}^\alpha_{pq} = -\frac{1}{2}g^{\alpha\beta}\partial_\beta\Phi_{pq} \tag{3.18}$$

We now compute the $\tilde{\omega}$ terms and we find, Greek indices raised with $g^{\alpha\beta}$, small Latin indices lowered with Φ_{mn},

$$\hat{\omega}^m_{\alpha\beta} = \tilde{\omega}^m_{\alpha\beta} = -\frac{1}{2}F^m_{\alpha\beta}, \quad \hat{\omega}^\alpha_{m\beta} = \hat{\omega}^\alpha_{\beta m} = -\frac{1}{2}F^\alpha{}_{\beta,m}. \tag{3.19}$$

$$\hat{\omega}^m_{pq} = \tilde{\omega}^m_{pq} = \omega^m_{pq} \tag{3.20}$$

where ω^m_{pq} is the Riemannian connection of the metric $\Phi_{mn}(x)\theta^m\theta^n$ induced by $\hat{g}$ on the orbit G_x, i.e. with $x = (x^\alpha)$ fixed; therefore

$$\hat{\omega}^m_{pq} \equiv \frac{1}{2}(C^m_{pq} - \Phi_{qr}\Phi^{ms}C^r_{ps} - \Phi_{pr}\Phi^{ms}C^r_{qs}).$$

[4] i.e. transforms by a change of trivialization of the bundle by a law of the type $F \to h^{-1}Fh$ (see CB-DM1 Vbis).

To write in compact form other connection coefficients we recall a notation and definition.

Definition 3.1 *One denotes by D the covariant derivative on sections of vector bundles over V associated with the G-principal fiber bundle $\hat{V} \to V$. This covariant derivative is taken in both the Riemannian connection of the metric g and the G-connection A. For instance it holds that:*

$$D_\gamma F^m_{\alpha\beta} \equiv \partial_\gamma F^m_{\alpha\beta} - \omega^\lambda_{\gamma\alpha} F^m_{\lambda\beta} - \omega^\lambda_{\gamma\beta} F^m_{\alpha\lambda} + C^m_{np} A^p_\gamma F^n_{\alpha\beta}. \tag{3.21}$$

and

$$D_\alpha \Phi_{mn} = \partial_\alpha \Phi_{mn} - C^q_{\alpha n} \Phi_{mq} - C^q_{\alpha m} \Phi_{qn}. \tag{3.22}$$

Using this definition we find that $\hat{\omega}^\alpha_{mn}$ is a vector on V with values in the tensor product $\mathcal{G}^* \otimes \mathcal{G}^*$, $\mathcal{G}^*$ dual of $\mathcal{G}$, given by

$$\hat{\omega}^\alpha_{mn} = -\frac{1}{2} g^{\alpha\beta} D_\beta \Phi_{mn}. \tag{3.23}$$

and

$$\hat{\omega}^m_{n\alpha} = \frac{1}{2} \Phi^{mq} \partial_\alpha \Phi_{nq} + \tilde{\omega}^m_{n\alpha} \tag{3.24}$$

with, using the values of the structure coefficients and their antisymmetry

$$\tilde{\omega}^m_{n\alpha} = \frac{1}{2} \{ \hat{C}^m_{n\alpha} - g_{\alpha\lambda} \Phi^{ms} \hat{C}^\lambda_{ns} - \Phi_{\ell n} \Phi^{pm} \hat{C}^\ell_{\alpha p} \} = -\frac{1}{2} \Phi^{mq} (\Phi_{sq} C^s_{np} + \Phi_{\ell n} C^\ell_{qp}) A^p_\alpha). \tag{3.25}$$

Hence

$$\hat{\omega}^m_{n\alpha} = \frac{1}{2} \Phi^{mq} D_\alpha \Phi_{nq} \tag{3.26}$$

and (we use the relation in Chapter 6 defining a torsionless connection)

$$\hat{\omega}^m_{\alpha n} = \hat{\omega}^m_{n\alpha} + \hat{C}^m_{\alpha n}. \tag{3.27}$$

4 Curvature tensor

We recall that the curvature tensor tensor of a pseudo-Riemannian metric $\hat{g} \equiv \hat{g}_{AB} \hat{\theta}^A \hat{\theta}^B$ in an arbitrary frame $\hat{\theta}^A$ is given in terms of the connection by

$$\hat{R}_{LM}{}^A{}_B \equiv \partial_L \hat{\omega}^A_{MB} - \partial_M \hat{\omega}^A_{LB} - \hat{\omega}^A_{CB} \hat{C}^C_{LM} + \hat{\omega}^A_{LC} \hat{\omega}^C_{MB} - \hat{\omega}^A_{MC} \hat{\omega}^C_{LB} \tag{4.1}$$

We find by computation, with $R_{\lambda\mu}{}^\alpha{}_\beta$ the curvature tensor of g,

$$\hat{R}_{\lambda\mu.}{}^\alpha{}_\beta \equiv R_{\lambda\mu}{}^\alpha{}_\beta + \frac{1}{2} F_{m,\beta}{}^\alpha F^m_{\lambda\mu} + \frac{1}{4} F_{m,\lambda}{}^\alpha F^m_{\beta\mu} - \frac{1}{4} F_{m,\mu}{}^\alpha F^m_{\beta\lambda} \tag{4.2}$$

and, with $R_{lm}{}^{p}{}_{q}$ the curvature tensor of the metric with constant coefficients (x is fixed) $\Phi_{mn}(x)\theta^m\theta^n$,

$$\hat{R}_{lm}{}^{p}{}_{q} \equiv R_{lm}{}^{p}{}_{q} - \frac{1}{4}g^{\gamma\alpha}\Phi^{rp}(D_\alpha\Phi_{mq}D_\gamma\Phi_{lr} - D_\gamma\Phi_{mr}D_\alpha\Phi_{lq}). \tag{4.3}$$

The definition (4.1) gives

$$\hat{R}_{\lambda m}{}^{\alpha}{}_{q} \equiv \partial_\lambda\hat{\omega}^{\alpha}_{mq} - \hat{\omega}^{\alpha}_{pq}C^{p}_{\lambda m} + \hat{\omega}^{\alpha}_{\lambda C}\hat{\omega}^{C}_{mq} - \hat{\omega}^{\alpha}_{mC}\hat{\omega}^{C}_{\lambda q}; \tag{4.4}$$

that is, with D the covariant derivative in the metric g on Greek indices and covariant derivative in the connection A on Latin indices

$$\hat{R}_{\lambda m}{}^{\alpha}{}_{q} \equiv -\frac{1}{2}D_\lambda D^\alpha\Phi_{mq} + \frac{1}{2}F_{p,\lambda}{}^{\alpha}\omega^{p}_{mq} + \frac{1}{4}F_{m,\gamma}{}^{\alpha}F_{q,\lambda}{}^{\gamma} - \frac{1}{4}\Phi^{pr}D^\alpha\Phi_{mp}D_\lambda\Phi_{qr} \tag{4.5}$$

On the other hand, with $f_{[AB]} = f_{AB} - f_{BA}$

$$\hat{R}_{lm}{}^{\alpha}{}_{q} \equiv \frac{1}{2}g^{\alpha\beta}D_\beta\Phi_{pq}c^{p}_{lm} + \frac{1}{4}F^{\gamma\alpha}_{[m}D_\gamma\Phi_{l]q} - \frac{1}{2}g^{\alpha\beta}D_\beta\Phi_{p[l}\omega^{p}_{m]q} + \frac{1}{4}F^{\gamma\alpha}_{[m}D_\gamma\Phi_{l]q} \tag{4.6}$$

and

$$\hat{R}_{\lambda\mu}{}^{\alpha}{}_{q} \equiv \frac{1}{2}D_{[\lambda}F^{\dots\dots\alpha}_{q,\mu]} + \frac{1}{2}g^{\alpha\beta}D_\beta\Phi_{pq}F^{p}_{\lambda\mu} + \frac{1}{4}F^{\dots\dots\alpha}_{p,[\lambda}\Phi^{pn}D_{\mu]}\Phi_{qn} \tag{4.7}$$

The other components of the curvature can be obtained from the symmetries properties of the Riemann tensor and the algebraic Bianchi identity.

Remark 4.1 *The transformation law (3.6) of Φ under a change of trivialization implies the identity*

$$D_\alpha D_\beta\Phi_{mq} \equiv D_\beta D_\alpha\Phi_{mq} + \rho_{\alpha\beta}{}^{p}{}_{m}\Phi_{pq} + \rho_{\alpha\beta}{}^{p}{}_{q}\Phi_{mp}, \tag{4.8}$$

where $\rho_{\alpha\beta}{}^{m}{}_{n}$ is the representative of the curvature of the gauge connection in the vector bundle associated to the given G-principal bundle by the adjoint representation; that is,

$$\rho_{\alpha\beta}{}^{m}{}_{n} \equiv C^{m}_{pn}F^{p}_{\alpha\beta}. \tag{4.9}$$

Therefore

$$\Phi^{mq}D_\alpha D_\beta\Phi_{mq} \equiv \Phi^{mq}D_\beta D_\alpha\Phi_{mq} + 2F^{r}_{\alpha\beta.}c^{p}_{rp} \tag{4.10}$$

In particular if the group is compact (up to the product by an Abelian group) we have the symmetry property

$$\Phi^{mq}D_\alpha D_\beta\Phi_{mq} = \Phi^{mq}D_\beta D_\alpha\Phi_{mq} \tag{4.11}$$

5 Ricci tensor and K–K equations

The components of the Ricci tensor of $(\hat{V}, \hat{g})$ result immediately from those of the curvature tensor. They simplify to

$$\hat{R}_{MB} := \hat{R}_{AM}{}^{A}{}_{B} \equiv \partial_A \hat{\omega}^A_{MB} - \partial_M \hat{\omega}^A_{AB} - \hat{\omega}^A_{CB}\hat{\omega}^C_{AM} + \hat{\omega}^A_{AC}\hat{\omega}^C_{MB} \tag{5.1}$$

After straightforward computation, using Remark 4.1 to make apparent the symmetry and simplify the expression of $\hat{R}_{\alpha\beta}$. We find

$$\hat{R}_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}F_{m,\beta}{}^{\mu}F^{m}_{\alpha\mu} - \frac{1}{4}\{\Phi^{mq}(D_\alpha D_\beta + D_\beta D_\alpha)\Phi_{mq} + D_\alpha \Phi^{mq} D_\beta \Phi_{mq}\}. \tag{5.2}$$

Definitions and computation also give

$$\hat{R}_{\alpha l} \equiv \frac{1}{2}D_\lambda F_{l,\alpha}{}^{\lambda} + \frac{1}{4}F_{l,\alpha}{}^{\lambda}\Phi^{np}D_\lambda \Phi_{np} + \frac{1}{2}c^p_{lq}\Phi^{nq}D_\alpha \Phi_{np} + \frac{1}{2}c^p_{pn}\Phi^{nq}D_\alpha \Phi_{lq} \tag{5.3}$$

$$\begin{aligned}\hat{R}_{mn} \equiv R_{mn} + \frac{1}{4}\{&F_{m,\alpha\beta}F^{\alpha\beta}_n + 2\Phi^{pq}D_\alpha \Phi_{mp}D^\alpha \Phi_{nq} \\ &- \Phi^{pq}D_\alpha \Phi_{pq}D^\alpha \Phi_{mn} - 2D_\alpha D^\alpha \Phi_{mn}\}\end{aligned} \tag{5.4}$$

with R_{mn} the Ricci tensor of the metric on G with constant coefficients (x fixed) $\Phi_{mn}\theta^m\theta^n$, that is,

$$R_{mn} \equiv -\omega^p_{qn}\omega^q_{pm} + \omega^p_{pq}\omega^q_{mn}, \tag{5.5}$$

where $\omega^p_{pq} = 0$ if the group is compact up to the product by an Abelian group.

Example. Let $G \equiv SU(2)$. The structure constants are then totally antisymmetric with

$$C^3_{12} = 1 \tag{5.6}$$

Suppose the metric Φ is diagonal, given by, with a_m some constants

$$\Phi \equiv \sum_{m=1,2,3} a_m (\theta^m)^2 \tag{5.7}$$

We find that the only non-zero connection coefficients of the metric φ are those with three different indices and

$$\omega^3_{12} = \frac{1}{2}(a_1 a_3^{-1} + a_2 a_3^{-1} - 1) = -\omega^3_{21} \tag{5.8}$$

We find that $R_{mn} = 0$ if $m \neq n$. Let us compute R_{11}; R_{22} and R_{33} are then obtained by circular permutation of indices. We find a result already obtained in Chapter 5 for Bianchi IX spacetimes.

$$R_{11} \equiv \frac{1}{2a_2a_3}\{a_1^2 - (a_2 - a_3)^2\} \tag{5.9}$$

6 Equations in conformal spacetime metric

As was already known from the five-dimensional theory, the Kaluza–Klein $\hat{d}$-dimensional vacuum equations $Ricci(\hat{g}) = 0$ take a form more familiar to physics if they are expressed with a metric $\tilde{g}$ on the basis V **conformal** to the original metric g. We recall that if

$$\tilde{g}_{\alpha\beta} \equiv \exp(2h) g_{\alpha\beta} \tag{6.1}$$

then (see Appendix VI),

$$R_{\alpha\beta} \equiv \tilde{R}_{\alpha\beta} + (n-1)(\nabla_\alpha \partial_\beta h - \partial_\alpha h \partial_\beta h) + g_{\alpha\beta}\{\nabla^\lambda \partial_\lambda h + (n-1)\partial^\lambda h \partial_\lambda h\}. \tag{6.2}$$

We suppose that the group G is compact (up to product by an Abelian group). Then c^q_{mp} is totally antisymmetric, in particular

$$C^p_{mp} = 0. \tag{6.3}$$

We have then the following lemma.

Lemma 6.1 *If the group G is compact (up to product by an Abelian group), then $\Phi^{mn} D_\beta \Phi_{mn}$ is a covariant vector on V given by:*

$$\Phi^{mn} D_\beta \Phi_{mn} \equiv \Phi^{mn} \partial_\beta \Phi_{mn} = \partial_\beta f, \tag{6.4}$$

with

$$f := log \det \Phi \tag{6.5}$$

where $\det \Phi$ is the determinant of the matrix (Φ_{mn}).

Proof By the definition of (Φ^{mn}) as the inverse matrix of (Φ_{mn}) and the definition of D it holds that

$$\Phi^{mn} D_\beta \Phi_{mn} \equiv \Phi^{mn} \partial_\beta \Phi_{mn} - 2 c^m_{pm} A^p_\beta. \tag{6.6}$$

The result of the lemma is then the consequence of the identity (6.3) satisfied by the structure constants of the considered groups, together with the formula giving the derivative of a determinant. □

We choose as conformal factor the function

$$\exp(2h) \equiv (det \Phi)^{\frac{1}{n-1}} \quad \text{i.e.} \quad h = \frac{1}{2(n-1)} f \tag{6.7}$$

and we prove the following lemma.

Lemma 6.2 *If $\tilde{g}_{\alpha\beta} \equiv \exp(2h) g_{\alpha\beta}$ with h given by (6.7), then the following identity holds:*

$$\hat{R}_{\alpha\beta} + \frac{g_{\alpha\beta}}{n-1} \Phi^{mn} \hat{R}_{mn} \equiv \tilde{R}_{\alpha\beta} - \rho_{\alpha\beta}. \tag{6.8}$$

where ρ is a tensor field on V given by

$$\rho_{\alpha\beta} \equiv \frac{1}{2}\left(F_{m,\beta}{}^{\mu}F^{m}_{\alpha\mu} - \frac{g_{\alpha\beta}}{2(n-1)}F_{m,\lambda\mu}F^{m,\lambda\mu}\right) \tag{6.9}$$

$$-\frac{1}{4}D_\alpha\Phi^{mq}D_\beta\Phi_{mq} + \frac{1}{4(n-1)}\partial_\alpha f\partial_\beta f - \frac{g_{\alpha\beta}}{(n-1)}\Phi^{mn}R_{mn}. \tag{6.10}$$

Proof Using the formulae (6.2) and (5.2) gives

$$\hat{R}_{\alpha\beta} \equiv \tilde{R}_{\alpha\beta} + (n-1)(\nabla_\alpha\partial_\beta h - \partial_\alpha h\partial_\beta h) + g_{\alpha\beta}\{\nabla^\lambda\partial_\lambda h - (n-1)\partial^\lambda h\partial_\lambda h\}.$$
$$-\frac{1}{2}F_{m,\beta}{}^{\mu}F^{m}_{\alpha\mu} - \frac{1}{2}\nabla_\alpha\partial_\beta f + \frac{1}{4}D_\alpha\Phi^{mq}D_\beta\Phi_{mq}. \tag{6.11}$$

The relation (6.7) between h and f implies the cancelling of the second derivatives $\nabla_\alpha\partial_\beta h$ and $\nabla_\alpha\partial_\beta f$, therefore

$$\hat{R}_{\alpha\beta} \equiv \tilde{R}_{\alpha\beta} + (n-1)(\nabla_\alpha\partial_\beta h - \partial_\alpha h\partial_\beta h) + g_{\alpha\beta}\{\nabla^\lambda\partial_\lambda h + (n-1)\partial^\lambda h\partial_\lambda h\}$$
$$-\frac{1}{2}\nabla_\beta\partial_\alpha f - \frac{1}{2}F_{m,\beta}{}^{\mu}F^{m}_{\alpha\mu} + \frac{1}{4}D_\alpha\Phi^{mq}D_\beta\Phi_{mq}. \tag{6.12}$$

We now compute $\Phi^{mn}\hat{R}_{mn}$, using (5.4) and the identity

$$\Phi^{mq}D_\alpha D_\beta\Phi_{mq} \equiv D_\alpha(\Phi^{mq}D_\beta\Phi_{mq}) - D_\alpha\Phi^{mq}D_\beta\Phi_{mq} = \nabla_\alpha\partial_\beta f - D_\alpha\Phi^{mq}D_\beta\Phi_{mq}. \tag{6.13}$$

We find

$$\Phi^{mn}\hat{R}_{mn} \equiv \Phi^{mn}R_{mn} + \frac{1}{4}F_{m,\alpha\beta}F^{m,\alpha\beta} - \frac{1}{4}\partial_\lambda f\partial^\lambda f - \frac{1}{2}\nabla^\lambda\partial_\lambda f. \tag{6.14}$$

We see that, when $f = 2(n-1)h$, the second derivatives of h and f cancel in the combination $\hat{R}_{\alpha\beta} + \frac{1}{n-1}g_{\alpha\beta}\Phi^{mn}\hat{R}_{mn}$, as well as the terms in $\partial_\lambda f\partial^\lambda f$ and $\partial_\lambda h\partial^\lambda h$. We thus obtain the formulae (6.8) and (6.9). □

Theorem 6.3 *If the isometry group G is compact (up to the product by an Abelian group), the $\hat{d}$-dimensional Einstein Kaluza–Klein equations* in vacuo, *$Ricci(\hat{g}) = 0$, can be written as an Einstein–Yang–Mills wave map type system on the $n+1$-manifold V.*

Proof Equations (5.4), $\hat{R}_{mn} = 0$, are wave equations in the metric g with quadratic non-linear terms in the first derivatives of Φ and in the curvature F.

Under the hypothesis made on the group G, Equations (5.3), $\hat{R}_{\alpha l} = 0$, reduce to

$$\hat{R}_{\alpha l} \equiv \frac{1}{2}D_\lambda F_{l,\alpha}{}^{\lambda} + \frac{1}{4}F_{l,\alpha}{}^{\lambda}\partial_\lambda f \equiv \frac{1}{2|\det\Phi|^{1:2}}D_\lambda(F_{l,\alpha}{}^{\lambda}|\det\Phi|^{1:2}) = 0 \tag{6.15}$$

which are equations for a weighted Yang–Mills field.

The previous lemma gives Einstein equations for the metric $\tilde{g}$ with sources quadratic in the first derivatives of Φ and a Yang–Mills type field. □

These results justify the name given to the system obtained for the unknowns $(\tilde{g}, \Phi, A)$. We will not pursue further here this general study. Applications are found in Chapter 14 and in the section "S^1 invariant Einstein universes" in Chapter 16.

RELATED PAPERS

This page intentionally left blank

Causality of classical supergravity

Yvonne Choquet-Bruhat
I.M.T.A, Université Paris 6.

Introduction

Supergravity (cf [1],[2]) has some of its origin in the difficulties raised by the classical Rarita Schwinger field equations for fields of spin 3/2 in a curved background. It also originates, and this point has been much more stressed, from two important aims. The first is to find a renormalizable theory of quantum gravity - such a theory, if it exists at all, has not been found yet, but supergravity improves matters in this respect, being at least one loop renormalizable. The other aim is to find a unified theory for boson and fermion fields.

In such a theory it should be possible to transform bosons into fermions and vice versa by a transformation - called a supersymmetry transformation - which would play the role that gauge transformations have assumed in the unification of various interactions by Yang-Mills fields, as in the Weinberg-Salam model. The simple, original, supergravity on 4-dimensional space-time M unifies gravity, a boson field of spin 2, with spin 3/2 field, through a Lagrangian - which is the action used for quantization. This Lagrangian can be proved to be invariant under the infinitesimal supersymmetry transformations

$$(1)\quad \delta\psi_\lambda = D_\lambda\, \varepsilon,$$

$$(2)\quad \delta e_a^{\ \lambda} = \bar{\varepsilon} v^\lambda\, \psi_a$$

where ψ_λ, the spin 3/2 field, is a spinor valued 1-form on M, and $e_a^{\ \lambda}$ a field of orthonormal frames, which defines the hyperbolic metric, the gravitational field, through the identity

$$g_{\lambda\mu} = e^a_{\ \mu} e^b_{\ \mu} \eta_{ab}$$

($e^a_{\ \lambda}$ inverse matrix of $e_a^{\ \lambda}$, η_{ab} Minkowski metric = diag (1,-1,-1,-1)) while ε is an arbitrarily given spinor field, D_λ its covariant derivative in an appropriately chosen connection, metric but with torsion.

(2) implies the infinitesimal transformation of $g_{\lambda\mu}$

$$\delta g_{\lambda\upsilon} = \bar{\varepsilon}(\delta_\lambda \psi_\mu + \gamma_\mu \psi_\lambda)$$

62

In fact the Lagrangian of supergravity is invariant under these infinitesimal transformations only under some restriction on ψ_λ, which is usually taken to be that ψ_λ should be a Majorana spinor, that is a self-conjugate field. In particle language that would mean that the particle it represents is identical to its antiparticle. Another possibility to obtain the existence of the infinitesimal supersymmetry transformations is to start from a slightly different real Lagrangian and choose ψ_λ to be a Weyl spinor of fixed helicity: the interpretation is then that the spin 3/2 particle of supergravity - the so called gravitino - is like the neutrino taking part in weak interactions, a parity violating particle.

In both cases a necessary hypothesis, more revolutionary for classical (I mean non quantum) fields, is that ψ_λ takes anticommuting values. Otherwise the infinitesimal supersymmetry invariance does not hold.

In fact this infinitesimal supersymmetry is not at all obvious on the Lagrangian, as it is for other invariances in physics, for instance, invariance by diffeomorphisms of the Einstein Lagrangian, or gauge invariance of the Yang-Mills Lagrangian. Indeed there is no simple formulation (if there is one at all) of finite (i.e. non infinitesimal) supersymmetry transformations. The existence of the infinitesimal supersymmetry relies on an identity which holds for the modified Rarita-Schwinger equation of Deser-Zumino (cf also [2]), due to their ingenious choice of this equation, under the hypothesis I have already stated on ψ_λ: that is ψ_λ is a Weyl (or Majorana) spinor and takes anticommuting values. This identity is also the integrability condition for the Einstein-Rarita-Schwinger (as modified by Deser-Zumino) system of partial differential equations deduced from the Lagrangian. Using this identity these equations can be proved to be a well posed, causal system.

A system of partial differential equations on a space-time M is said to be well posed if when one has initial data for the unknown at some given time, for instance $t = 0$, satisfying eventually some constraints imposed by the condition that the initial data must satisfy the equations at $t = 0$, then the system has a solution, at least for some time $t < T$ (of course the initial data must belong to some functional space, depending on the problem at hand). The system is said to be causal if the solution at some point depends only on the initial data in the past of this point, the past being determined by the light cone of the metric $g_{\lambda\mu}$.

The fact that the Rarita-Schwinger equation for a spin 3/2 field in the presence of an electromagnetic potential is not a causal system

had already been noticed by Fierz and Pauli, and the associated so called "acausal propagation" has given rise to various works (cf [10], [12]) - in fact I would rather say in this case that there is no propagation since the constraints are not preserved through evolution. The system is not only acausal, but not well posed.

In this paper I shall first recall why the R.S. equation is causal in Minkowski space time and is not so in a curved space time. Then I shall review briefly the fundamentals of classical, Grassmann valued supergravity (d = 4, N = 1). I shall write the identity which result from the invariance of the Lagrangian (for a general Dirac ψ_λ field) through Lorentz transformations in the tangent space and diffeomorphism of the space time, and prove the supersymmetry identity, for a Weyl, odd valued in the Grassmann algebra, ψ_λ field. I shall then prove, by an appropriate choice of gauges, that the system is indeed well posed and causal.

1 - Rarita-Schwinger equations in Minkowski space time.

Physical considerations had led Rarita and Schwinger to propose that, in Minkowski space-time, a spin 3/2 field, represnted by a vector-spinor ψ_λ, should satisfy the differential system[3]:

$$(1\text{-}1) \qquad A^\lambda \equiv \varepsilon^{\lambda\mu\nu\rho}_{0123}\gamma_\mu \partial_\nu \psi_\rho = 0$$

together with the linear equation

$$(1\text{-}2) \qquad \chi \equiv \gamma^\lambda \psi_\lambda = 0$$

Rarita-Schwinger argued that (1-2) can be considered as a constraint on the spin 3/2 field characterizing "pure spin states". A constraint, for a system of partial differential equations on space time, is a relation between the unknown which, at each time depends only on the Cauchy data relative to the system. They are usually a consequence of the existence of an invariance group for the system, and their counterpart is the existence of identities, which should, for an acceptable theory, insure their preservation under time evolution.

We study from this point of view the Cauchy data the system (1-1) (1-2), at t = 0, are [5] $\psi_\alpha(0,.)$. They must satisfy the constraint

$$(1\text{-}3) \qquad A^\circ \equiv \varepsilon^{ijk}_{123}\gamma_i \partial_j \psi_k \qquad \text{for } t = 0$$

64

(1-4) $\gamma^{\lambda}\psi_{\lambda} = 0$ for $t = 0$

On the other hand the equation (1-1) satisfies obviously the identity

(1-5) $\partial_{\lambda}A^{\lambda} \equiv 0.$

It is then possible to show that the system (1-1)(1-2) is well posed and causal.

a) One has the identity[6]

(1-6) $2A_{\alpha} + \xi\gamma^{\mu}(\partial_{\mu}\psi_{\alpha} - \partial_{\alpha}\psi_{\mu}) - \gamma_{\alpha}\gamma_{\lambda}A^{\lambda} \equiv 0$

therefore, if we set

(1-7) $f_{\alpha} \equiv \gamma^{\mu}\partial_{\mu}\psi_{\alpha}$

We have

(1-8) $2A_{\alpha} + \xi f_{\alpha} - \xi\partial_{\alpha}\chi + \gamma_{\alpha}\gamma_{\lambda}A^{\lambda} \equiv 0$

thus $(\gamma^{\alpha}\gamma_{\alpha} = -4, \gamma^{\alpha}\xi = -\xi\gamma^{\alpha})$

(1-9) $2\gamma^{\alpha}A_{\alpha} + \xi\gamma^{\alpha}f_{\alpha} - \xi\gamma^{\alpha}\partial_{\alpha}\chi = 0$

and

(1-10) $2A_{\alpha} - \tfrac{1}{2}\xi\gamma_{\alpha}f_{\lambda} + \tfrac{1}{2}\xi\gamma^{\lambda}\gamma_{\alpha}\partial_{\lambda}\chi \equiv 0$

It results from (1-5) that every solution of (1-1),(1-2) satisfies the hyperbolic system of classical Dirac equations

(1-11) $f_{\alpha} \equiv \gamma^{\mu}\partial_{\mu}\psi_{\alpha} = 0$

b) Conversely the identity (1-3) implies that every solution of $f_{\alpha} = 0$ satisfies

(1-12) $\gamma^{\lambda}\gamma^{\alpha}\partial_{\alpha}\partial_{\lambda}\chi \equiv -\eta^{\alpha\lambda}\partial_{\alpha}\partial_{\lambda}\chi = 0$

If we take Cauchy data $\psi_{\lambda}(0,.)$ which satisfy $A_{o} = 0$ and $\chi = 0$ for $t = 0$ we have also (cf(1-6)) $\partial_{o}\chi = 0$ for $t = 0$. Equation (1-7)

then implies that $\chi = 0$ for all t: for such Cauchy data (1-6) implies (1-1) and (1-2). We have proved:

Theorem: The free Rarita-Schwinger equations (1-1),(1-2) in Minkowski space time are a well posed causal physical system, with the constraints (1-2),(1-3).

2 - Rarita-Schwinger equations in curved space time.

The most natural generalization of (1-1),(1-2) in a curved space time, four dimensional manifold V with an hyperbolic metric g, is

$$(2\text{-}1) \qquad A^{\lambda} \equiv \eta^{\lambda\mu\nu\rho}\gamma_{\lambda}\nabla_{\nu}\psi_{\rho} = 0$$

$$(2\text{-}2) \qquad \gamma^{\lambda}\psi_{\lambda} = 0$$

where the $\eta^{\lambda\mu\nu\rho}$ are the contravariant components[7] of the volume 4-form, and γ_{λ} the covariant vector-spinor-cospinor defining the Clifford algebra: if e^{a}_{λ} is an orthonormal frame relative to g, and e^{λ}_{a} the inverse matrix

$$e^{a}_{\lambda}e^{b}_{\mu}\eta_{ab} = g_{\lambda\mu}, \quad e^{\lambda}_{a}e^{\mu}_{b}\eta^{ab} = g^{\lambda\mu}$$

we have[8]

$$\gamma_{\lambda} = e^{a}_{\lambda}\gamma_{a}$$

where the γ_{a} are a set of standard Dirac matrices (cf Note (1)).

∇ is the Riemannian covariant derivative: for the covariant vector-spinor ψ_{ρ} we have[9]:

$$(2\text{-}3) \qquad \nabla_{\nu}\psi_{\rho} = \partial_{\nu}\psi_{\rho} - \Gamma^{\alpha}_{\nu\rho}\psi_{\alpha} + \sigma_{\nu}\psi_{\rho}$$

where σ_{ν} is the representant in the spin algebra of the Riemannian connexion

$$(2\text{-}4) \qquad \sigma_{\nu} = -\frac{1}{4}\omega_{\nu}{}^{a}{}_{b}\gamma_{a}\gamma^{b}$$

We not that

$$(2\text{-}5) \qquad \nabla_{\lambda}\gamma_{\mu} \equiv \partial_{\lambda}\gamma_{\mu} - \Gamma^{\alpha}_{\lambda\mu}\gamma_{\alpha} + \sigma_{\lambda}\gamma_{\mu} - \sigma_{\mu}\gamma_{\lambda} \equiv 0$$

66

We deduce from the Ricci identity for vector-spinors (cf [9]), if $R_{\lambda\nu,\rho\sigma}$ is the curvature of g:

$$(2\text{-}6) \qquad \nabla_\lambda A^\lambda \equiv \eta^{\lambda\mu\nu\rho}\gamma_\mu\nabla_\lambda\nabla_\nu\psi_\rho \equiv \tfrac{1}{2}\eta^{\lambda\mu\nu\rho}\gamma_\mu\,(R_{\lambda\nu\rho}{}^\sigma\psi_\sigma - \tfrac{1}{4}R_{\lambda\nu,ab}\gamma^{ab}\psi_\rho)$$

The first term vanishes by the algebraic Bianchi identity for a riemannian connexion, while the second gives (cf Note (6))

$$(2\text{-}7) \qquad \nabla_\lambda A^\lambda \equiv -\tfrac{1}{2}G^\rho_\lambda\gamma^\lambda\psi_\rho \ , \ G^\rho_\lambda \equiv R^\rho_\lambda - \tfrac{1}{2}\,g^\rho_\lambda\,R$$

$G_{\lambda\rho}$, $R_{\lambda\rho}$, R are respectively the Einstein tensor, the Ricci tensor and the scalar curvature of g. The identity (2-7) gives an "integrability condition" for (2-1): any solution of (2-1) satisfies the algebraic relation

$$(2\text{-}8) \qquad G^\rho_\lambda\gamma^\lambda\psi_\rho = 0,$$

any solution of (2-1) and (2-2) satisfies

$$(2\text{-}9) \qquad R^\rho_\lambda\gamma^\lambda\psi^\rho = 0.$$

The condition (2-9) reduces to (2-2) for an Einstein space ($R_{\alpha\beta} = \lambda g_{\alpha\beta}$) but for other space times it restricts severely the generality of the spin 3/2 field.

Indeed A_α and

$$(2\text{-}10) \qquad f_\alpha \equiv \gamma^\mu\nabla_\mu\psi_\alpha$$

satisfy identities analogous to the ones obtained in §1 in the flat case. If we denote by χ the spinor

$$(2\text{-}11) \qquad \chi \equiv \gamma^\lambda\psi_\lambda$$

we still have

$$(2\text{-}12) \qquad 2A_\alpha - \tfrac{1}{2}\xi\gamma^\lambda\gamma_\alpha f_\lambda + \tfrac{1}{2}\xi\gamma^\lambda\gamma_\alpha\nabla_\lambda\chi \equiv 0$$

Therefore, every solution of (2-1),(2-2) satisfies the hyperbolic system of Dirac equations, with characteristics the light cone of the metric $g_{\lambda\mu}$

(2-13) $\quad f_\lambda \equiv \gamma^\mu \nabla_\mu \psi_\lambda = 0$

But conversely a solution of (2-13) satisfies only

$$2\nabla_\alpha A^\alpha + \tfrac{1}{2}\xi\gamma^\lambda\gamma^\alpha\nabla_\alpha\nabla_\lambda\chi \equiv - G_\lambda{}^\rho\gamma^\lambda\psi_\rho - \tfrac{1}{2}\xi(\nabla^\alpha\nabla_\alpha\chi + \tfrac{1}{4}R\,\chi) = 0$$

and the initial conditions

$$\chi = 0,\ A_o = 0 \quad \text{for} \quad t = 0$$

are not preserved in general by evolution through the hyperbolic system[10] (2-13). The system is not well posed.

3 - Supergravity

The idea of supergravity (Deser-Zumino [3], Ferrara, Freedman, Van Nieuwenhuisen [6]) is to modify the Rarita-Schwinger equation (2-1) in such a way that (2-7) is replaced by an identity, which will insure the existence of solutions for the spin 3/2 field with causal propagation. This identity will be valid modulo the equation itself, and a generalisation of Einstein equations, with source the spin 3/2 field. The counterpart of this identity will be an infinitesimal invariance group for the coupled system, the so-called supersymmetry transformation which mixes the fermionic spin 3/2 field (gravitino) and the bosonic spin 2 field (graviton). It is natural to look for such a coherent system in the framework of Einstein-Cartan theory, adapted to the presence of a momentum density. It turns out that it is possible to find a well posed causal system for an Einstein-Cartan spin 2 field coupled with a Rarita-Schwinger spin 3/2, but only by assuming that the spinor fields are anticommuting, that is, their components in a frame are not numbers, but take their values in a Grassmann algebra $\mathcal{G}$. If we denote by ζ^I $I = 1,\ldots,N$ (eventually $N = +\infty$), a set of generators of $\mathcal{G}$, an element of $\mathcal{G}$ is a sum (a formal series if $N = \infty$)

$$(3\text{-}1) \qquad f = \sum_{p=0}^{N} f(p)$$

with $f(p)$ an element of order p of $\mathcal{G}$, that is

$$f(p) = \frac{1}{p!}\, f_{I_1\ldots I_p}\zeta^{I_1}\ldots\zeta^{I_p}$$

where $f_{I_1\ldots I_p}$ are numbers (real or complex).

The algebra is defined by the usual sum and product laws (of formal

series if $N = \infty$) with the anticommutation rule

$$\zeta^I \zeta^J = - \zeta^J \zeta^I$$

An even [resp. odd] element of (3-1) contains only even [resp. odd] terms. Odd elements anticommute, even elements commute between themselves and with odd elements. The term of order zero of an element is a number, it is called the body while the remaining part is called the soul (DeWitt, Van Nieuwenhuisen, West [5]).

The inverse of the element f with non vanishing body $f(o)$ and soul f_S is:

$$f^{-1} \equiv (f(o) + f_S)^{-1} = f(o)^{-1} \sum_{n=0}^{N} (-1)^n (f(o) f_S)^n$$

The definitions of differential geometry carry on for $\mathcal{G}$ -valued objects on a differentiable manifold, by imposing to their Grassmann valued components the equivalence relations which define tensors, spinors, connections through their usual numerical valued representant in different charts and frames.

The derivatives of $\mathcal{G}$ -valued tensors in a $\mathcal{G}$ -valued connection are again sums (or formal series) obtained by replacing in the classical formulas the connection coefficients by their $\mathcal{G}$ -valued expression. The connection in the following will be even-valued, so there will be no ordering problem. The standard laws for differentiation of tensor products or contractions are valid. Of course one must be careful to respect the order of non commuting elements.

The unknown usually considered as fundamental in Supergravity are the tetrad field e_λ^a, even-valued in $\mathcal{G}_\lambda$, with invertible body[11] $e_\lambda^a(o)$ hence with an inverse in $\mathcal{G}$, denoted e_a^λ, and the spin 3/2 field ψ_α, odd-valued in $\mathcal{G}$. The gravitational spin 2 field, is the even-valued in $\mathcal{G}$, symmetric 2 tensor

$$(3\text{-}2) \qquad g_{\lambda\mu} = \eta^{ab} e_\lambda^a e_\mu^b$$

where η^{ab} is the (numeric !) Minkowski metric. The body of $g_{\lambda\mu}$ is an ordinary hyperbolic metric

$$g_{\lambda\ \mu}^{(o)} = \eta^{ab} e_\lambda^a(o) e_\mu^a(o)$$

$g_{\lambda\mu}$ is invariant under a Lorentz transformation $e_\lambda^{a'} \rightsquigarrow L_a^{a'} e_\lambda^a$.

The equations of supergravity derive formally from the variational

principle[12] with real, $\mathcal{U}$-valued Lagrangian

$$(3\text{-}3) \qquad \int (\hat{R} - i(\bar{\psi}_\lambda \xi A^\lambda + \bar{A}^\lambda \xi \psi_\lambda)) d\mu(g)$$

where $d\mu(g)$ is the volume element of the metric, that is, in coordinates:

$$d\mu(g) = \det(e_\lambda{}^a)\, d^4x.$$

We denote by $\bar{\psi}$ the Dirac adjoint of the spinor ψ, that is

$$\bar{\psi} = i\tilde{\psi}\gamma_o$$

where $\tilde{\psi}$ is the Hermitian conjugate of ψ

$$\tilde{\psi} = (\psi^T)^\ddagger, \quad \ddagger \text{ denotes complex conjugate.}$$

For anticommuting spinors we have

$$(3\text{-}4a) \qquad (\bar{\psi}\varphi)^\ddagger = -\bar{\varphi}\psi, \quad (\bar{\psi}\xi\varphi)^\ddagger = -(\bar{\varphi}\xi\psi)$$

$$(3\text{-}4b) \qquad (\bar{\psi}\gamma_a\varphi)^\ddagger = \bar{\varphi}\gamma_a\psi, \quad (\bar{\psi}\xi\gamma_a\varphi)^\ddagger = -(\bar{\varphi}\xi\gamma_a\psi)$$

A^λ is the vector-spinor

$$(3\text{-}5) \qquad A^\lambda \equiv \eta^{\lambda\mu\nu\rho}\gamma_\mu D_\nu \psi_\rho$$

where the derivative D_ν is the Riemannian covariant derivative in tensor spaces[13], and a metric covariant derivative, but with torsion, in spinor spaces. This subtlety in the definition of covariant derivatives is an essential point in supergravity. We denote by ω the Riemannian connection [resp. $\hat{\omega}$ the metric connection with torsion] it is $\mathcal{U}$-even-valued, defined in the natural frame by Christoffel symbols $\Gamma^\lambda_{\alpha\beta}$ of $g_{\lambda\mu}$ [resp. the connection coefficients $\hat{\Gamma}^\lambda_{\alpha\beta}$]. The difference of the two connections is called the contorsion tensor $C^\lambda_{\alpha\beta}$:

$$(3\text{-}6) \qquad \hat{\Gamma}^\lambda_{\alpha\ \beta} = \Gamma^\lambda_{\alpha\ \beta} + C^\lambda_{\alpha\ \beta}$$

If $\hat{\Gamma}$ is, as Γ is, a metric connection we have

$$(3\text{-}7) \qquad C_{\alpha\gamma\beta} + C_{\alpha\beta\gamma} = 0$$

70

From (3-7) and the definiton of the torsion tensor

$$S_{\alpha\ \gamma}^{\ \beta} = C_{\alpha\ \gamma}^{\ \beta} - C_{\gamma\ \alpha}^{\ \beta}$$

we deduce

(3-8) $$C_{\alpha\ \gamma}^{\ \beta} = \tfrac{1}{2}(S_{\alpha\ \gamma}^{\ \beta} - S_{\alpha\gamma}^{\ \ \beta} + S^{\beta}_{\ \alpha\gamma})$$

$\hat{R}$ is the scalar curvature of the connection $\hat{\omega}$, whose curvature tensor is $\hat{R}_{\lambda\mu,\nu\rho}$. It is convenient to rewrite $\mathscr{L}$ as follows, where $\{\ \}$ denotes the imaginary part:

(3-9) $$\mathscr{L} = \int (R_{\lambda\mu}^{\ \ ab} e_a^\lambda e_b^\mu + 2\{\bar{\psi}_\lambda \xi A^\lambda\})\ \det\ (e_\lambda^a) d^4x$$

We obtain the equations of supergravity by varying independantly ψ_λ, $\bar{\psi}_\lambda, e_a^\lambda$ and $\hat{\omega}_\lambda^{\ ab}$, connection 1-form in the orthonormal frame[14] (we know that $\hat{\omega}_\lambda^{\ ab} = -\hat{\omega}_\lambda^{\ ba}$, since $\hat{\omega}$ is metric): the equations express then vanishing of the coefficients of these variations. We shall obtain thus, $\mathscr{G}$-valued equations. In the process of writing these equations we shall perform as usual integration by parts and discard Riemannian divergences of vector fields[15]. The result is:

(3-10) $$\delta\mathscr{L} \equiv \int 4\{\delta\bar{\psi}_\lambda \xi B^\lambda\} + \delta\hat{\omega}_\nu^{\ ab} D^\nu_{ab} + 2\delta\ e_a^\lambda \hat{\Sigma}_\lambda^{\ a}) d\mu(g)$$

with

(3-11) $$B^\lambda \equiv A^\lambda + \tfrac{1}{2}\eta^{\lambda\mu\nu\rho} D_\nu \gamma_\xi \psi_\rho,\ D_\nu \gamma_\mu = C^\alpha_{\nu\mu}\gamma_\alpha$$

(3-12) $$D^\nu_{\ ab} \equiv \{\bar{\psi}_a \gamma^\nu \psi_b\} - \{\bar{\psi}_{[a} \chi e^\nu_{\ b]}\} + C^\mu_{\mu\lambda} e^\lambda_{[a} e^\nu_{\ b]} + e\lambda_a e^\mu_b S_\lambda^{\ \ \nu}{}_\mu$$

with

$$\chi = \gamma^a \psi_a,\ f_{[ab]} = f_{ab} - f_{ba}$$

(3-13) $$\hat{\Sigma}_\lambda^{\ a} \equiv \hat{G}_\lambda^{\ a} - T_\lambda^{\ a}$$

with (one uses $\delta e_\lambda^a = -e_\lambda^{\ b} e_\mu^a \delta e_b^\mu$)

(3-14) $$T_\lambda^a \equiv \eta^{a\beta\nu\rho}\{\bar{\psi}_\beta \gamma_\lambda \xi D_\nu \psi_\rho\}$$

Equating to zero B^λ, $D^\nu_{\ ab}, \hat{\Sigma}_\lambda^{\ a}$ gives the equations of classical

supergravity. We first show that $D^{\nu}{}_{ab} = 0$ determines S, in terms of ψ. Indeed, since $e^{c}{}_{\nu}$ is an even element of $\mathcal{G}$:

$$(3\text{-}15)\quad D_a{}^c{}_b \equiv e^c_\nu D^\nu_{ab} \equiv \{\bar\psi_a \gamma^c \psi_b\} - \{\bar\psi_{[a}\chi\delta^c_{b]}\} - S^c_{ab} + C^m_{m[b}\,\delta^c_{a]} = 0$$

We deduce from (3-8) and the antisymmetry of S

$$C_m{}^m{}_a = S_m{}^m{}_a = -\,S_a{}^m{}_m$$

Thus $D_a{}^b{}_b = 0$ gives:

$$(3\text{-}16)\quad C_b{}^b{}_a = \{\bar\psi_a \chi\}$$

and (3-15) is equivalent to the definition of S as the even element of $\mathcal{G}$:

$$(3\text{-}17)\quad S_a{}^c{}_b = -\,\{\bar\psi_a \gamma^c \psi_b\}$$

We shall always take (3-17) as a definition for the torsion of the metric connection $\hat\omega$. The first variation δ thus reduces to

$$(3\text{-}18)\quad \delta\mathcal{A} \equiv \int (4\{\delta\bar\psi_\lambda \xi B^\lambda\} + 2\delta e_a^\lambda\, \hat\Sigma_\lambda{}^a)\, d\mu(g)$$

and, the equations of supergravity reduce to, for instance in the natural frame

$$(3\text{-}19)\quad B^\lambda \equiv \eta^{\lambda\mu\nu\rho}(\gamma_\mu D_\nu \psi_\rho + \tfrac{1}{2} C_\nu{}^\alpha{}_\mu \gamma_\alpha \psi_\rho) = 0$$

$$(3\text{-}20)\quad G_\lambda{}^\alpha = T_\lambda{}^\alpha,\ T_\lambda{}^\alpha \equiv \eta^{\alpha\beta\nu\rho}\{\bar\psi_\beta \gamma_\lambda \xi D_\nu \psi_\rho\}$$

4 - Invariances and identities

The Lagrangian (3-3) is invariant under Lorentz transformations of the orthonormal frame, and under a change of local coordinates. Indeed if we replace the $\mathcal{G}$-valued unknown $e_\lambda{}^a$ and ψ_λ by

$$e'_\lambda{}^{a'} = L_a{}^{a'} e_\lambda^a \quad \text{and} \quad \psi'_\lambda = \Lambda\psi_\lambda$$

where L is the Lorentz transformation associated to the spin transformation[16] Λ, the $\mathcal{G}$-valued integrand in (3-3) is unchanged, thus $\delta\mathcal{A}$, given by (3-18) since S is given by (3-17), is zero whenever

$\delta e_\lambda{}^a = U^a{}_b e_\lambda{}^b$, $\delta\psi_\lambda = -\tfrac{1}{4}U^{ab}\gamma_a\gamma_b\psi_\lambda$, with $U^{ab} = -U^{ba}$

a generator of a Lorentz transformation (in the tangent space $T_x V$, depending on $x \in V$). We have therefore the $\mathcal{G}$-valued identities[17]

$$(4\text{-}1) \qquad \hat{\Sigma}_{[ab]} + \{\psi_\lambda \xi \gamma_{ab} B^\lambda\} \equiv 0$$

For solutions of $B^\lambda = 0$ the equations (3-20) are thus equivalent to the symmetrized equations

$$\Sigma_{\alpha\beta} \equiv \hat{G}_{(\alpha\beta)} - T_{(\alpha\beta)} = 0$$

which can also be written, by denoting $G_{\alpha\beta}$ the Einstein tensor of the metric g

$$(4\text{-}2) \qquad \Sigma_{\alpha\beta} \equiv G_{\alpha\beta} - \Theta_{\alpha\beta} = 0$$

where $\Theta_{\alpha\beta}$ is a polynomial in ψ_λ, $D_\nu\psi_\lambda$ and their Dirac adjoints, and $e_\lambda{}^a$. If we take into account the identity (4-1) the invariance of $\mathcal{L}$ under coordinate changes leads to the identity (∇ is the covariant Riemannian derivative)

$$(4\text{-}3) \qquad \nabla_\mu \hat{\Sigma}_\lambda{}^\mu - 2\{(\nabla_\alpha\bar{\psi}_\lambda - \nabla_\lambda\bar{\psi}_\alpha)\xi B^\alpha + \bar{\psi}_\lambda \xi \nabla_\alpha B^\alpha\} \equiv 0$$

Supersymmetry

To insure the coherence of the theory we must check that $D_\lambda B^\lambda = 0$ does not imply new algebraic relations between the unknowns. We shall show that, if we restrict the space of spinors appropriately, we have $D_\lambda B^\lambda = 0$, at least "on shell", i.e. modulo the other equations. From the definition (3-11) of B^λ we deduce:

$$D_\lambda B^\lambda = \eta^{\lambda\mu\nu\rho}\gamma_\mu D_\lambda D_\nu \psi_\rho + \tfrac{1}{2}\eta^{\lambda\mu\nu\rho}(D_\lambda D_\nu \gamma_\mu)\psi_\rho + \tfrac{1}{2}\eta^{\lambda\mu\nu\rho}(D_\lambda\gamma_\mu)D_\nu\psi_\rho$$

According to the definition of D the Ricci identity gives[18]

$$(D_\lambda D_\nu - D_\nu D_\lambda)\psi_\rho \equiv R_{\lambda\nu\rho}{}^\alpha \psi_\alpha - \tfrac{1}{4}\hat{R}_{\lambda\nu,ab}\gamma^{ab}\psi_\rho$$

and, using some Dirac algebra[19]

$$(D_\lambda D_\nu - D_\nu D_\lambda)_\mu \equiv (R_{\lambda\nu,\mu\alpha} - \hat{R}_{\lambda\nu,\mu\alpha})\gamma^\alpha$$

73

Thus, by the Bianchi identity for Riemannian curvature

$$D_\lambda B^\lambda = -\frac{1}{8}\eta^{\lambda\mu\nu\rho}\hat{R}_{\lambda\nu,\alpha\beta}(\gamma_\mu\sigma^{\alpha\beta} + 2\delta_\mu^\alpha\gamma^\beta)\psi_\rho + \tfrac{1}{4}\eta^{\lambda\mu\nu\rho}S^\alpha_{\lambda\mu}\gamma_\alpha D_\nu\psi_\rho$$

Using again Dirac algebra we find[20]

$$D_\lambda B^\lambda \equiv -\tfrac{1}{2}\hat{G}_\lambda{}^\rho\xi\gamma^\lambda\psi_\rho + \tfrac{1}{4}\eta^{\lambda\mu\nu\rho}S^\alpha_{\lambda\mu}\gamma_\alpha D_\nu\psi_\rho$$

which we write

$$D_\lambda B^\lambda \equiv -\tfrac{1}{2}\hat{\Sigma}_\lambda{}^\rho\xi\gamma^\lambda\psi_\rho + C \tag{4-4}$$

where C is the $\mathcal{U}$-valued spinor field defined by

$$C = \tfrac{1}{4}\eta^{\lambda\mu\nu\rho}S_\lambda{}^\alpha{}_\mu\gamma_\alpha D_\nu\psi_\rho - \tfrac{1}{2}T_\lambda{}^\rho\xi\gamma^\lambda\psi_\rho \tag{4-5}$$

$T_\gamma{}^\rho$ is given by (3-14), thus:

$$T_\lambda{}^\alpha\xi\gamma^\lambda\psi_\alpha = -\frac{i}{2}\eta^{\alpha\beta\nu\rho}(\bar{\psi}_\beta\gamma_\lambda\xi D_\nu\psi_\rho + D_\nu\,\bar{\psi}_\rho\gamma_\lambda\xi\psi_\beta)\xi\gamma^\lambda\psi_\alpha \quad .$$

To simplify the expression of C one uses the Fierz identity for anticommuting spinors[21]

$$(\bar{\psi}\mathcal{P})(\bar{\chi}\lambda) = \tfrac{1}{4}\sum_J \varepsilon_J(\bar{\psi}\Gamma_J\lambda)(\bar{\chi}\Gamma^J\mathcal{P})$$

where Γ_J, $J = 1,\ldots,16$ are the generators I, ξ, γ_a, $\xi\gamma_a$, γ_{ab} of the Dirac algebra and $\Gamma^J = (I,\xi,\gamma^a,\xi\gamma^a,\gamma^{ab})$, $\varepsilon_1 = -1$ and $\varepsilon_J = 1$ if $J \neq 1$.

Using this identity we obtain

$$(\bar{\psi}_\beta\gamma_\lambda\xi D_\nu\psi_\rho)\xi\gamma^\lambda\psi_\alpha = \tfrac{1}{4}\sum_J \varepsilon_J(\bar{\psi}_\beta\gamma_\lambda\xi\Gamma_J\xi\gamma^\lambda\psi_\alpha)\Gamma^J D_\nu\psi_\rho \tag{4-6}$$

and

$$(4\text{-}7) \qquad (D_\nu \bar{\psi}_\rho \gamma_\lambda \xi \psi_\beta)\xi\gamma^\lambda\psi_\alpha = \tfrac{1}{4} \sum_J \varepsilon_J (D_\nu \bar{\psi}_\rho \gamma_\lambda \xi \Gamma_J \xi \gamma^\lambda \psi_\alpha)\Gamma^J\psi_\beta$$

C does not vanish for general Dirac spinors, but we shall show that it vanishes for Weyl spinors[22]. This means that the gravitino is, like the neutrino, taking part in weak interactions, a parity violating particle.

A Weyl spinor is one which takes its values in a given eigenspace of the matrix ξ[23]. For instance we impose

$$\xi\psi_\lambda = i\psi_\lambda, \text{ thus } \quad \bar{\psi}_\lambda \xi = -i\bar{\psi}_\lambda$$

For a pair φ , ψ of Weyl spinors one has the identities

$$\bar{\varphi}\psi = 0 \ , \quad \bar{\varphi}\xi\psi = 0$$

Since $\gamma_\lambda\gamma^\lambda = -4, \gamma_\lambda\sigma^{ab}\gamma^\lambda = 0$ the identity (4-6) reduces, for Weyl spinors to

$$\tfrac{1}{2}(\bar{\psi}_\beta\gamma_\lambda\gamma_a\gamma^\lambda\psi_\alpha)\gamma^a D_\nu\psi_\rho = (\bar{\psi}_\beta\gamma_\lambda\psi_\alpha)\gamma^\lambda D_\nu\psi_\rho$$

while (4-7) gives an expression symmetric in α and β:

$$(D_\nu\bar{\psi}_\rho\gamma_\lambda\psi_\beta)\gamma^\lambda\psi_\alpha = (D_\nu\bar{\psi}_\rho\gamma_\lambda\psi_\alpha)\gamma^\lambda\psi_\beta$$

Therefore, for a Weyl spinor

$$\begin{aligned} T_\lambda{}^\alpha\xi\gamma^\lambda\psi_\alpha &= -\frac{i}{2}\eta^{\alpha\beta\nu\rho}\,(\bar{\psi}_\beta\gamma_\lambda\psi_\alpha)\gamma^\lambda D_\nu\psi_\rho \\ &= +\tfrac{1}{2}\eta^{\alpha\beta\nu\rho}\{\psi_\beta\gamma_\lambda\psi_\alpha\}\gamma^\lambda D_\nu\psi_\rho \\ &= \tfrac{1}{2}\eta^{\alpha\beta\nu\rho}S_\alpha{}^\lambda{}_\beta\gamma^\lambda D_\nu\psi_\rho \end{aligned}$$

The identity (4-3) for a Weyl spinor reduces to

$$(4\text{-}8) \qquad D_\lambda B^\lambda = \frac{i}{2}\hat{\Sigma}_\lambda{}^\rho\gamma^\lambda\psi_\rho$$

which implies

$$(4\text{-}8\text{bis}) \qquad D_\lambda\bar{B}^\lambda = \frac{i}{2}\hat{\Sigma}^\rho_\lambda\,\bar{\psi}_\rho\gamma^\lambda$$

For a Weyl spinor the identity (4-3) and (4-1) imply

75

(4-9) $\nabla_\lambda \Sigma_\mu{}^\lambda = 0 \mod \Sigma_\lambda{}^\rho, B^\alpha$

and we have also

(4-10) $D_\lambda B^\lambda = 0 \mod \Sigma_\lambda{}^\rho, B^\alpha$

Remark For a Weyl spinor we have

$$\eta^{\lambda\mu\nu\rho} C_\nu{}^\alpha{}_\mu \gamma_\alpha \psi_\rho = \tfrac{1}{2}\eta^{\lambda\mu\nu\rho} S_\nu{}^\alpha{}_\mu \gamma_\alpha \psi_\rho \equiv 0$$

since

$$\eta^{\lambda\mu\nu\rho} S_\nu{}^\alpha{}_\mu \gamma_\alpha \psi_\rho = i\eta^{\lambda\mu\nu\rho} (\bar\psi_\nu \gamma^\alpha \psi_\mu)\gamma_\alpha \psi_\rho$$

and $(\bar\psi_\nu \gamma^\alpha \psi_\mu)(\gamma_\alpha \psi_\rho)$ is symmetrical in μ and ρ by the Fierz identity for a Weyl anticommuting spinor. Thus in this case

$$B^\lambda \equiv A^\lambda$$

From the identity (4-8) one deduces a new infinitesimal invariance of the theory, namely the infinitesimal invariance of the Lagrangian by the "supersymmetry": for a Weyl spinor (3-18) reduces to

$$\delta\mathcal{L} = 2\int (-\delta\bar\psi_\lambda B^\lambda + \bar B^\lambda \delta\psi_\lambda + \delta e_a^\lambda \hat\Sigma_\lambda{}^a)\, d\mu(g)$$

If we take

(4-11) $\delta\psi_\lambda = D_\lambda \varepsilon, \quad \delta\bar\psi_\lambda = D_\lambda \bar\varepsilon$

we see by using the identities[24] (4-8) and (4-8bis) that $\delta\mathcal{L} \equiv 0$ if

(4-12) $\delta e_a^\lambda = \{\bar\varepsilon \gamma^\lambda \psi_a\}$

(4-11), (4-12) is the infinitesimal supersymmetry.

5 - Cauchy problem

To have a manageable system[25] we consider as unknown ψ_λ, odd valued in $\mathcal{U}$, $g_{\lambda\mu}$ and $e_\lambda{}^a$, even valued in $\mathcal{U}$ and linked by the relation:

(5-1) $g_{\lambda\mu} e^\lambda{}_a e_b^\mu = \eta_{ab}$

The equations to satisfy are (cf (4-2))

$$G_{\alpha\beta} = \Phi_{\alpha\beta} \tag{5-2}$$

$$B^{\lambda} = 0 \tag{5-3}$$

We have identically, by the Dirac algebra (cf Note (6)) for an arbitrary spinor

$$2A_{\lambda} + \gamma_{\lambda}(\gamma_{\mu}A^{\mu}) + \xi\gamma^{\mu}(D_{\mu}\psi_{\lambda} - D_{\lambda}\psi_{\mu}) \equiv 0$$

thus

$$2B_{\lambda} + \gamma_{\lambda}(\gamma_{\mu}B^{\mu}) + \xi f_{\lambda} - \xi D_{\lambda}\chi \equiv 0 \tag{5-4}$$

with, for a Weyl spinor[26] (cf remark §4)

$$f_{\lambda} \equiv \gamma^{\mu}D_{\mu}\psi_{\lambda} + C_{\lambda}{}^{\rho}{}_{\alpha}\gamma^{\alpha}\psi_{\rho} \tag{5-5}$$

From (5-4) we deduce

$$-2\gamma_{\lambda}B^{\lambda} - \xi\gamma^{\lambda}f_{\lambda} + \xi\gamma^{\lambda}D_{\lambda}\chi = 0 \tag{5-6}$$

and therefore

$$2B_{\lambda} - \tfrac{1}{2}\xi\gamma^{\mu}\gamma_{\lambda}f_{\mu} + \tfrac{1}{2}\xi\gamma^{\mu}\gamma_{\lambda}D_{\mu}\chi = 0 \tag{5-7}$$

If ψ_{λ} is a Weyl spinor satisfying $f_{\lambda} = 0$ we have $D_{\lambda}B^{\lambda} = 0$ modulo $\hat{\Sigma}_{\lambda}{}^{\rho}$ thus also

$$D^{\lambda}D_{\lambda}\chi = 0 \quad \text{modulo} \quad \chi,\ D_{\mu}\chi,\ \Sigma_{\lambda}{}^{\rho}. \tag{5-8}$$

Let S be a 3-dimensional, C^{∞}, manifold

The Cauchy data on S are

a) A $\mathcal{U}$- odd valued vector field φ^{i} with values in the eigenspace $S_{+} \subset \mathbb{C}^{4}$ of ξ.

b) A $\mathcal{U}$-even valued moving frame ε_{u}^{i} $(i,u = 1,2,3)$

We define a $\mathcal{U}$-even valued metric γ by ($\varepsilon_{i}{}^{u}$ inverse matrix of $\varepsilon_{u}{}^{i}$)

$$\gamma_{ij} = \varepsilon_{i}{}^{u}\varepsilon_{j}{}^{v}\delta_{uv}\ ,\ \delta_{u} = \mathrm{diag}(1,1,1) \tag{5-9}$$

and we give on S a $\mathcal{G}$-even valued symmetric 2-tensor k_{ij}. We imbedd S in $S \times \mathbb{R}$ and identify S and $S_o = S \times \{0\}$. We always choose coordinates x^i, x^o, adapted to the product structure. We complete the Cauchy data by the following conditions (to be considered as initial gauge conditions).

c) We define the Lorentzian moving frame of $S \times \mathbb{R}$ at each point of S by the vectors of a) and a vector ε_o^α tangent to $\mathbb{R}$, i.e. $\varepsilon_o^i = 0$, and for instance $\varepsilon_o^o = 1$. The chosen standard representation of Dirac matrices γ_a and ξ are, by hypothesis, linked to these moving frames (and their Lorentz transforms).

The $\mathcal{G}$-valued hyperbolic metric is, on S

$$(5\text{-}10) \qquad g_{\lambda\mu}dx^\lambda dx^\mu = N_o^2(dx^o)^2 - \gamma_{ij}dx^i dx^j$$

with $N_o = \varepsilon_o^o$, for instance $N_o = 1$.

We complete the Cauchy data on S_o by (initial gauge conditions)

$$(5\text{-}11) \qquad \partial_o N|_{S_o} = \operatorname{tr} k, \quad \partial_o g_{io}|_{S_o} = 0$$

(note that $\partial_o g_{ij}|_{S_o} = -2\, k_{ij}$)

d) We complete the data of the Weyl vector spinor φ on S_o by the initial gauge condition

$$(5\text{-}12) \qquad \chi|_{S_o} \equiv \gamma_\lambda \varphi^\lambda = 0$$

Constraints

The initial data must satisfy on S the equation, which depends only on these data

$$(5\text{-}13) \qquad B^o = 0 \quad \text{on} \quad S_o$$

It is well known on the other hand that $G^{o\alpha}|_{S_o}$ depend only on γ_{ij}, k_{ij}, N. The equations $B_\lambda|_{S_o} = 0$ and $\chi|_{S_o} = 0$ imply $f_i|_{S_o} = 0$, and thus determine $\partial_o \psi_i|_{S_o}$. We choose $\partial_o \psi_o|_{S_o}$ such that $D_o\chi|_{S_o} = 0$ (then $B_\lambda|_{S_o} = 0$ implies also $f_o|_{S_o} = 0$). The derivatives $\partial_o \psi_\lambda|_{S_o}$ are thus determined in terms of the Cauchy data, and these data must satisfy the constraints

78

(5-14) $\Sigma^{\alpha o} \equiv G^{\alpha o} - \mathfrak{m}^{\alpha o} = 0$

Evolution

We consider the system (5-1),(5-2),(5-3) in the gauges

(5-15) $\gamma^{\lambda}\psi_{\lambda} = 0$

(5-16) $N^2 \equiv g_{oo} = ((\det \gamma_{ij})^{-1} \det g_{ij})^{\frac{1}{2}}$, $g_{io} = 0$

which are compatible with the initial data.

We replace (5-3) by

(5-17) $f_{\lambda} \equiv \gamma^{\mu} D_{\mu}\psi_{\lambda} + C_{\lambda}{}^{\rho}{}_{\alpha}\gamma^{\alpha}\psi_{\rho}$

and (5-2) by the system (method introduced in [2] for the classical Einstein system - one could also use harmonic coordinates cf [13]). $\bar{\nabla}$ denotes the Riemannian covariant derivative in the space metric g_{ij}

(5-18a) $\partial_o g_{ij} = -2NK_{ij}$, $K_{ij} = P_{ij} - 2g_{ij}P$

$$\partial_o \Sigma^{ij} + N^2(\bar{\nabla}^i \Sigma^{jo} + \bar{\nabla}^j \Sigma^{io} - g^{ij}\bar{\nabla}_h \Sigma^{ho}) = 0$$

which is of the type

(5-19) $(N^{-2}\partial_o^2 - \bar{\nabla}^h\bar{\nabla}_h)\, P^{ij} + F^{ij} = 0$, $P^{ij} = K^{ij} - g^{ij}K^h_h$

where F^{ij}, symmetric in i and j is a polynomial in g_{hk}, g^{hk}, $\partial_{\alpha}g_{hk}$, $\partial^2_{\alpha\beta}\, g_{hk}$, $P^{ij}, \partial_{\alpha}P^{ij}, \psi_{\alpha}, \partial_{\alpha}\psi_{\lambda}, \partial^2_{\alpha\beta}\psi_{\lambda}$, e^{λ}_a.

We replace the equation (5-18) by[27]

(5-18b) $\partial_o e_i{}^a = -e^{aj}NK_{ij}$,

which implies (5-8) since K_{ij} is symmetric.

Remark The equation (5-18b) also implies

$$e_{ai}\partial_o e_j{}^a - e_{aj}\partial_o e_i{}^a = 0,$$

an analogue of the 0(3,1) gauge condition of Bao-Isenberg-Yasskin (cf [14]). If the unknown were scalar valued the system (5-17), (5-18a), (5-19) would be a classical hyperbolic system, with propagation determined by the isotropic cone of the metric $g_{\lambda\mu}$. The identities (4-9) and (4-10) would insure the preservation of the gauge conditions, and constraints, under time evolution. Indeed, if we choose for the equations the Leray weights s(5-17) = 0, s(5-18a) = 0, s(5-19) = -2 and for the unknown ψ_λ, $e_i^{\ a}$ (with i,a = 1,2,3) a,d P^{ij} the respective Leray weights $t(\psi_\lambda) = 1$, $t(e_i^{\ a}) = 1$, $t(P^{ij}) = 0$ we find that our system has a determinant of the type (where the sign ⊛ denotes some uncomputed expressions, and the dots in the

$$J = \begin{pmatrix} \gamma^\lambda \xi_\lambda & \circledast & \circledast \\ 0 & \xi_0 \cdots & \circledast \\ 0 & 0 & g^{\lambda\mu}\xi_\lambda\xi_\lambda \end{pmatrix}$$

diagonal means that the expressions are repeated, 4,9, or 10 times. The value of this determinant is

$$J = (g^{\lambda\mu}\xi_\lambda\xi_\mu)^{18}(\xi_0)^9$$

The system is hyperbolic non strict in the sense of Leray-Ohya (cf criterion in [15]). If the unknown were scalar valued the Cauchy problem would have a solution in a Gevrey class - C^∞ functions with some restriction on the derivatives - if the data were also in an appropriate Gevrey class. The system would thus be well posed in these functional spaces, and causal, since the domain of dependence of the solution is determined by the light cone, the time line being interior to this cone. However, since the fields are $\mathcal{G}$ -valued, the equations demultiply into a recursive system of equations, written for each subspace of $\mathcal{G}$ of a given order. Each of these systems is strictly hyperbolic, and only the first one, the numerical valued Einstein equations, is non linear. Indeed since ψ is odd, $S^\lambda_{\alpha\beta}$ and $\Theta_{\alpha\beta}$ are even, or $\mathcal{G}$ -degree at least 2. The body of the system reduces to the classical Einstein equations in empty space for the body of the metric $g_{\lambda\mu}$:

(5-20) $$G_{\alpha\beta}(g_{\lambda\mu}(o)) = 0$$

and the gauge conditions for the body of e_a^λ

(5-21) $\quad e_i^{\ o} = e_o^{\ i} = 0 \ , \ e_{[i}^{\ \ a}\, {}_o e_{j]\, a} = 0$

The terms of order 1 reduce to

(5-22) $\quad B^\lambda(1) \equiv \eta^{\lambda\mu\nu\rho}(o)\gamma_\mu(o)\nabla_\nu(o)\psi_\rho(1) = 0$

The solution of the Cauchy problem for (5-20) is well known, it is a classical causal system with constraints ($G^{\alpha o}(o) = 0$). The gauge condition (5-21) permits a causal determination of $e_a^{\ \lambda}(0)$.

Then $g_{\alpha\beta}(o)$, $e_a^{\ \lambda}(o)$ are determined, (5-22) is also a classical causal system with constraint ($B^o(1) = 0$), since the integrability condition found in §2 is satisfied by (5-20).

More generally the terms of the $\hbar$ expansion for $g_{\lambda\mu}$, $e_a^{\ \lambda}$ and ψ_λ are determined by recursive classical causal systems with constraints whose characteristic cones are the isotropic cone of the body $g_{\lambda\mu}(o)$; we see this as follows:

We deduce from the $\mathfrak{g}$-valued identity (5-4) when $B^\lambda(p) = 0$, $f_\lambda(p) = 0$, $\chi(p) = 0$, $\Sigma_{\lambda\mu}(p+1) = 0$ for $p \leq 2n-1$ that

(5-23) $\quad 2B_\lambda(2n+1) + \gamma_\lambda(o)\gamma_\mu(0)\, B^\mu(2n+1) + \xi f_\lambda(2n+1) - \xi D_\lambda(0)\chi(2n+1) \equiv 0$

and

(5-24) $\quad D_\lambda(0)\, B^\lambda(2n+1) \equiv 0 \mod B^\lambda(2n+1)$

The equation $f_\lambda(2n+1) = 0$ is of type:

(5-25) $\quad f_\lambda(2n+1) \equiv \gamma^\mu(o)\nabla_\mu(o)\psi_\lambda(2n+1) + F_\lambda = 0$

where F_λ is known if $\psi_\lambda(p)$, $e_\lambda^{\ a}(p+1)$ are known for $p \leq 2n-1$. The identities (5-23) and (5-24) show that a solution of (5-25) (causal, linear, hyperbolic system) satisfies $B_\lambda(2n+1)$ if the Cauchy data satisfy $B^o(2n+1) = 0$ and $\chi(2n+1) = 0$ (which are consequences of $B^o = 0$ and $\chi = 0$ on S).

An analogous reasoning shows that $g_{\lambda\mu}(2n+2)$, $e_\lambda^{\ a}(2n+2)$ are determined by a causal, linear hyperbolic system, when $e_\lambda^{\ a}(p-1)$ and $\psi_\lambda(p)$ are known for $p \leq 2n+1$, if they satisfy the constraints $\Sigma^{\alpha o}(2n+2) = 0$.

Remark We have shown the existence of a solution in chosen gauges, with data satisfying initial gauge conditions. For the Einstein equations it is well known that an arbitrary, globally hyperbolic solution is isometric, in a neighborhood of a Cauchy surface, to a solution in the standard gauges (harmonic, or 3+1 as we choose here). The same property is not known of the Rarita-Schwinger gauge.

81

Acknowledgement - I thank Jim Isenberg and Phil Yasskin for stimulating discussions.

Notes

(1) Our theory is a variant of the usual one which uses Majorana spinors, i.e. fields invariant under charge conjugaison.

(2) It seems impossible to realize this property with classical, commuting, spinors.

(3) Notations: $\varepsilon^{\lambda\mu\nu\rho}_{0123}$ totally antisymmetric Kronecker symbol, γ_μ Dirac matrices, ∂_ν partial derivative $\partial/\partial x^\nu$. Greek indices range from zero to 3, latin indices from 1 to 3. An hyperbolic metric g, for instance Minkowski η, has signature (+,-,-,-). The Dirac matrices satisfy $\gamma_\alpha\gamma_\beta + \gamma_\beta\gamma_\alpha = -2\,\eta_{\alpha\beta}I$, γ_i is Hermitian and γ_o anti-Hermitian.

(4) Physical uniqueness is also an interesting problem to consider.

(5) We denote by (t,x), $x^o = t$, $x = (x^i)$ a point of space time.

(6) As a consequence of the identity

$$\gamma_\alpha\gamma_{\lambda\mu} = -\eta_{\alpha\lambda\mu\beta}\gamma^\beta\xi - g_{\alpha\lambda}\gamma_\mu + g_{\alpha\mu}\gamma_\lambda$$

$$\gamma_{\lambda\mu} = \tfrac{1}{2}(\gamma_\lambda\gamma_\mu - \gamma_\mu\gamma_\lambda),\ \xi = \gamma^0\gamma^1\gamma^2\gamma^3$$

(7) $\eta^{\lambda\mu\nu\rho} = |\det g|^{-\frac{1}{2}}\varepsilon^{\lambda\mu\nu\rho}_{0123}$

(8) We shall, as usual, denote by Latin letters from the beginning of the alphabet components relative to an orthonormal frame.

(9) We denote by $\Gamma^\nu_{\lambda\mu}$ [resp. ω^c_{ab}] the connection coefficients in the natural frame [resp. the orthonormal frame]. They are related by the identity, with $\omega_\lambda{}^a{}_b = e^c_\lambda\omega_c{}^a{}_b$:

$$\nabla_\lambda e^a{}_\mu \equiv \partial_\lambda e^a_\mu - \Gamma_\lambda{}^\nu{}_\mu e^a_\nu + \omega_\lambda{}^a{}_b e^b_\mu = 0$$

(10) This phenomenon is also called "acausal propagation" (cf [12], [10]).

(11) $\det e_\lambda{}^a(0) \neq 0$ implies that $\det e_\lambda{}^a$ is invertible in $\mathcal{U}$. The inverse matrix $e_a{}^\lambda$ is defined by the usual algebraic formulas, because $e_\lambda{}^a$ is even.

(12) We do not suppose that the spinors are Majorana spinors (i.e. invariant under charge conjugation, i.e. real for a real representation of the Dirac matrices).

(13) We can also say that A is the spinor valued 1 form

$$A = i\,{}^*\gamma D\psi \text{ , i.e. } A_\lambda = \gamma^{\mu\,*}(D\psi)_{\lambda\mu}$$

where ${}^*D\psi$ is the spinor valued 2-form, adjoint of the spinor valued 2-form $D\psi$ defined by

82

$$D\psi = d\psi + \sigma\hat{}\psi$$

where d is the usual differential of a form with values in a given vector space, and σ the spin-connection one-form, with values in the linear endomorphisms of spinors (cf [9], [18]).

(14) We have $\hat{\omega}_\lambda{}^{ab} = e_\lambda{}^c \hat{\omega}_c{}^{ab}$. The connection coefficients in orthonormal and natural frames are linked by

$$\hat{\nabla}_\lambda e_\mu{}^a \equiv \partial_\lambda e_\mu{}^a + \hat{\omega}_\lambda{}^a{}_b e_\mu{}^b - \hat{\Gamma}^\nu_{\lambda\mu}\, e_\nu{}^a \equiv 0$$

Note that we have

$$\hat{\nabla}_\lambda \gamma_\mu = \nabla_\lambda \gamma_\mu = 0$$

but

$$D_\lambda \gamma_\mu = C_\lambda{}^\alpha{}_\mu \gamma_\alpha$$

(15) Such a divergence multiplied by $d\mu(g)$ is, in local coordinates, an ordinary divergence on an open set of $\mathbb{R}^4$, and its contribution to $\mathcal{L}$ is an integral on the boundary where the variations are supposed to vanish.

(16) Both numerical valued!

(17) Which generalize identities valid in ordinary Einstein-Cartan theories. These identities can also be checked directly.

(18) Since D is metric on tensor indices no torsion teams appear.

(19) $\gamma^{ab}\gamma^c - \gamma^c\gamma^{ab} \equiv 2\,\eta^{ca}\gamma^b - 2\,\eta^{cb}\gamma^a$

(20) $\gamma_c\gamma_{ab} \equiv -\eta_{cabd}\gamma^d\,\xi - \eta_{ca}\gamma_b + \eta_{cb}\gamma_a$

(21) Consequence of the identity between standard Dirac matrices

$$\delta^A_B\delta^C_D \equiv \tfrac{1}{4}\sum_J \varepsilon_J \Gamma_J{}^A{}_D \Gamma^J{}_B{}^C, \quad \varepsilon_1 = 1,\ \varepsilon_J = -1,\ J \neq 1$$

(22) The identity $D_\lambda B^\lambda = 0$ mod. $\hat{\Sigma}^\rho_\lambda$ is proved in the literature for Majorana spinors, starting from a variant of the Lagrangian.

(23) ξ has 2 eigenvalues, i and - i, and 2 eigenspaces of complex dimension 2.

(24) Recall that on tensor fields D reduces to the Riemannian covariant derivative where divergence integrates to zero.

(25) For an alternate analysis cf [16], [17].

(26) An analogous expression with an f_λ having the same principal terms is valid for a general spinor.

(27) One can also choose to determine the evolution of the tetrad field by its parallel transport along time lines (cf [13]).

83

REFERENCES

[1] Y. Choquet-Bruhat - Solutions globales des équations de Maxwell-Dirac-Klein-Gordon. Rendi Conti dell Circolo Mat. Palermo, Série II 31 (1982) 267-288.

[2] Y. Choquet-Bruhat et T. Ruggeri - Hyperbolicité du système 3+1 des équations d'Einstein. C.R. Ac. Sc. Paris 284, I (6) p. 225(1982)

[3] S. Deser and B. Zumino - Consistent Supergravity. Phys. Letters 62 B n°3 (1976), 335-337.

[4] S. Deser and B. Zumino - Broken supersymmetry and supergravity. Phys. Rev. Letters 38 n°25, 1433 - 1436 (1977).

[5] B.DeWitt, P. Van Nieuwenhuisen and P. West - Supermanifolds and supersymmetry, to appear.

[6] D. Z. Freedman, P. Van Nieuwenhuisen and S. Ferrara - Progress towards a theory of Supergravity. Phys. Rev. D, 13 n°12(1976) 3214-3218.

[7] F.W. Hehl, P. van der Heyde, G.D. Kerlick and J.M. Nester - General Realitivity with spin and torsion. Foundations and prospects. Rev. Mod. Phys. 48 395 (1976).

[8] C. Latrémolière - Sur les équations d'ordre supérieur du champ gravitationnel. C.R. Ac. Sc. Paris 258, 1170 (1964). Thèse Université Paris "Sur les champs libres de spin élevé en Relativité Générale" (1970).

[9] A. Lichnerowicz - Champ de Dirac, champ du neutrino et transformation CPT su un expace temps courbe. Ann. I.H.P. I, n°3, (1964) p. 233-290.

[10] J. Madore - The characteristic surfaces of a classical spin 3/2 field in an einstein background. Phys. Letters 55 B n°2 (1975).

[11] A. Trautman - Fiber bundles, gauge fields and gravitation in "General Relativity and Gravitation" A. Held ed. Plenum 1980.

[12] G. Velo and D. Zwanziger. Propagation and quantization of RARITA-SCHWINGER waves in an Essential Electromagnetic Potential. Phys. Rev. D 186 (1969) p. 1337.

[13] Y. Choquet-Bruhat - The Cauchy Problem in Classical Supergravity. Letters in Maths. Phys. to appear.

[14] D. Bao, J. Isenberg, P. Yasskin. The dynamics of the Einstein-Dirac system. Ann. of Physics, Part 1 (1982), Part 2 (1983).

[15] Y. Choquet-Bruhat. Diagonalisation des systèmes quasilinéaires et hyperbolicité non stricte. J. Maths. Pures et appl. 45 (1966) 371-386.

[16] D. Bao, Some aspects in the dynamic supergravity. Ph D thesis, Berkeley 1983.

[17] D. Bao, Y. Choquet-Bruhat, J. Isenberg, P. Yasskin, in preparation.

[18] J. Isenberg, J. Nester and R. Skinner, GR 8 à Sgax (1977).

[19] D. Bao, J. Isenberg, P. Yasskin, Is supergravity well-posed? Proc. of the third Grossmann meeting (1982).

Questions

1. Q. Can one use the Rarita-Schwinger gauge condition $\gamma^u \psi_u = 0$ when working with the Grassmann variable formulation of supergravity? Are there no inconsistencies?

   A. Yes. It is fine.

2. Q. Do you have the infamous decoupling phenomenon when you use the Grassmann formulation?

   A. Yes.

3. Q. Why must the Grassmann algebra be based on an infinite dimen-

sional algebra?

A. To permit an unlimited number of independent gauge transformations, all of which take values in the Grassmann algebra.

4. Q. What happens if you consider a Maxwell field coupled to a spin 3/2 field?

A. Presumably the coupling can be done consistently and in a supersymmetric way. And presumably the resulting theory, when formulated using a Grassmann algebra, is well-posed and decouples.

GRAVITATION WITH GAUSS BONNET TERMS

Yvonne Choquet-Bruhat

Abstract We study general properties of the partial differential equations for generalized gravity arising from the Lovelock Lagrangian.

I - Introduction

The low energy limit of supersymmetric string action leads to an effective Lagrangian which contains, in addition to the usual scalar curvature, polynomial terms in the Riemann curvature tensor; these terms are identically zero if the base manifold is of dimension four. The corresponding field equations are second order partial differential equations for the metric tensor if the Lagrangian was formed with the so called Gauss-Bonnet combination , that is if the polynomial of order p was of the type

$$L(p)= \mathcal{E}^{\lambda_1\lambda_2\cdots\lambda_{2p-1}\lambda_{2p}}_{\mu_1\mu_2\cdots\mu_{2p-1}\mu_{2p}} R_{\lambda_1\lambda_2}{}^{\mu_1\mu_2} \cdots R_{\lambda_{2p-1}\lambda_{2p}}{}^{\mu_{2p-1}\mu_{2p}}$$

which, if the manifold has dimension 2p, corresponds to a closed form which represents the Euler class of the manifold (up to a constant factor). The equations are non linear even for the second derivatives if the Lagrangian contains a Riemann curvature term of order $p > 1$. Various interesting results are already known for these field equations. First it is obvious that at flat space their first variation is identical whith the variation of Einstein equations. It has also been proved (Boulware and Deser) that they have the same plane wave solutions than Einstein equations. The same authors have constructed spherically symmetric solutions and studied their stability. Kerner has used these equations in a five dimensionnal Kaluza Klein context as a model for non linear electromagnetism. In the case p=2 the characteristics have been studied by Aragone . Other results pertinent to particular cases have also been obtained.

54

In this article we give some general properties of the generic solutions of the system of non linear partial differential equations, deduced from the Lagrangian with Gauss-Bonnet terms : we show the splitting of the equations like in ordinary gravity , between constraints and evolution, and show that at least in the analytic case the evolution preserves the constraints. We determine the wave fronts ; they are no more tangent to the light cone and not even to a second order cone , neither probably in general to a convex cone even if the polynomial terms in Riemann curvature are coupled with the ordinary Einstein tensor through multiplication by a small constant : the wave cone of the Einstein equations consists in fact of D copies of the light cone of the metric , and by perturbation it becomes a cone of order 2D which may be non convex and non real. Existence of solutions of the Cauchy problem are not known without analyticity hypothesis on the data.

In a second part we give some general results about high frequency waves.

II - Equations

Let V be a d-dimensional C^∞ manifold, with a metric g of hyperbolic signature (-, +, +, ..+). This metric is said to represent gravitation with Gauss-Bonnet terms if it satisfies the system of partial differential equations

$$A_\alpha^\beta \equiv K_o\, \delta_\alpha^\beta + \sum_{p=1}^{\infty} K_p\, \mathcal{E}_\alpha^{\beta\ \lambda_1 \cdots \lambda_{2p}}_{\mu_1 \cdots \mu_{2p}}\, R_{\lambda_1\lambda_2}{}^{\mu_1\mu_2} \cdots R_{\lambda_{2p-1}\lambda_{2p}}{}^{\mu_{2p-1}\mu_{2p}} = 0 \qquad (2\text{-}1)$$

where the K_p are constants, the $\mathcal{E}_\alpha^{\beta\ \lambda_1 \cdots \lambda_{2p}}_{\mu_1 \cdots \mu_{2p}}$ the completely antisymmetric Kronecker tensor, and $R_{\alpha\beta}{}^{\lambda\mu}$ the Riemann curvature tensor of g. The sum is indeed finite : all terms with $2p + 1 > d$ are identically zero; for $d = 4$ the only non zero terms are for $p = 1$. For arbitrary d the term in $p = 1$ is proportional to the Einstein tensor :

55

$$\mathcal{E}_{\alpha\ \mu_1\mu_2}^{\beta\ \lambda_1\lambda_2}\ R_{\lambda_1\lambda_2}{}^{\mu_1\mu_2} \equiv -\ 4\ \left(R_\alpha^\beta - \frac{1}{2}\ g_\alpha^\beta\ R\right) \equiv -\ 4\ S_\alpha^\beta \quad .$$

while the term with p = 0 is the so called cosmological term.

To make appear a coupling constant, eventually small, between the usual Einstein equations in dimension d with possibly a non zero cosmological constant, and the new Gauss-Bonnet terms, we rescale (2-1) to write it under the form

$$A_\alpha^\beta \equiv S_\alpha^\beta + \Lambda\ g_\alpha^\beta + \varkappa\ B_\alpha^\beta = 0 \qquad (2\text{-}2)$$

with, P < d/2 being some positive integer

$$B_\alpha^\beta \equiv \sum_{p=1}^{P}\ k_p\ \mathcal{E}_{\alpha\ \mu_1\ldots\mu_{2p}}^{\beta\ \lambda_1\ldots\lambda_{2p}}\ R_{\lambda_1\lambda_2}{}^{\mu_1\mu_2}\ldots R_{\lambda_{2p-1}\lambda_{2p}}{}^{\mu_{2p-1}\mu_{2p}}$$

We have, by the symmetries of the Riemann tensor

$$A_{\alpha\beta} \equiv A_{\beta\alpha} \quad ,$$

the equations (2-2) are, in local coordinates, a system of d(d+1)/2 second order partial differential equations for the d(d+1)/2 unknowns $g_{\alpha\beta}$.

These equations are invariant by diffeomorphisms of V : in fact the lagrangian itself is invariant by these diffeomorphisms, which implies, for any choice of the constants $\varkappa$ and k_p, the d identities

$$\nabla_\alpha\ A_\beta^\alpha \equiv 0 \quad ; \qquad (2\text{-}3)$$

as can be checked directly.

III - Cauchy problem. Constraints.

The equations (2-2) are, like the equations of ordinary General Relativity both an under determined system and an over determined one : their

characteristic determinant is identically zero at most of rank $d(d+1)/2 - d = d(d-1)/2$ due to (2-3) on the one hand, and on the other hand the unknowns and their first derivatives cannot be given arbitrarily on a d-1 dimensional submanifold of V : the Cauchy data must satisfy constraints.

To make a geometric analysis of the Cauchy problem we use, like in ordinary gravity, a (d-1)+1 decomposition of the metric and, in the case considered here, of the full riemann tensor.

Let $U = S \times I$ be a local slicing of an open set $U \subset V$ by d-1 dimensional space like manifolds $S_t = S \times \{t\}$. The metric reads, in adapted coordinates $x^\circ = t$, x^i coordinates on S :

$$ds^2 = - \alpha^2 (dx^\circ)^2 + g_{ij} (dx^i + \beta^i \, dx^\circ)(dx^j + \beta^j \, dx^\circ) \quad .$$

If the shift β is zero (choice always possible) a simple calculation gives the identities

$$R_{ijh}{}^k \equiv \bar{R}_{ijh}{}^k + K_j{}^k \, K_{ih} - K_i{}^k \, K_{hj} \tag{3-1-a}$$

$$R_{ijh}{}^\circ \equiv - \frac{1}{\alpha} \left(\bar{\nabla}_j \, K_{ih} - \bar{\nabla}_i \, K_{jh} \right) \tag{3-1-b}$$

$$R_{ioh}{}^\circ \equiv - \frac{1}{\alpha} \partial_\circ \, K_{ih} + K_{ij} \, K_h{}^j - \frac{1}{\alpha} \bar{\nabla}_i \, \partial_h \alpha \tag{3-1-c}$$

where $\bar{\nabla}$ and $\bar{R}_{ijh}{}^k$ are the riemannian covariant derivative and curvature tensor of the metric $\bar{g} = (g_{ij})$ induced on S_t by $g = (g_{\alpha\beta})$, and $K = (K_{ij})$ is the extrinsic curvature of S_t, that is

$$K_{ij} = - \frac{1}{2\alpha} \partial_\circ \, g_{ij} \quad . \tag{3-2}$$

We remark on the formulas (3-1) that the derivative $\partial^2_{\circ\circ} \, \alpha$ appears nowhere and that the derivative $\partial_\circ \, K_{ij}$, therefore the derivative $\partial^2_{\circ\circ} \, g_{ij}$, appears only in $R_{ioh}{}^\circ$. As a consequence the quantities $A^\circ_\circ$ and A°_i are determined on a slice S_t by the values on S_t of the first derivatives of the

metric, they give constraints on the Cauchy data, namely in the coordinates we have adopted :

$$A^{\circ}_{o} \equiv S^{\circ}_{o} + \Lambda\, g^{\circ}_{o} + \varkappa \sum_{p=2}^{P} k_p\, \mathcal{E}^{i_1 \cdots i_{2p}}_{j_1 \cdots j_{2p}}\, R_{i_1 i_2}{}^{j_1 j_2} \cdots R_{i_{2p-1} i_{2p}}{}^{j_{2p-1} j_{2p}}$$

$$A^{\circ}_{i} \equiv S^{\circ}_{i} - \varkappa \sum_{p=2}^{P} 2\, p\, k_p\, \mathcal{E}^{i_1 \cdots i_{2p}}_{j_2 \cdots j_{2p}}\, R_{i_1 i_2}{}^{\circ j_2}\, R_{i_3 i_4}{}^{j_3 j_4} \cdots R_{i_{2p-1} i_{2p}}{}^{j_{2p-1} j_{2p}}$$

We know from Einstein's equations that

$$S^{\circ}_{o} \equiv -\frac{1}{2}\left(\overline{R} - K_i{}^j\, K_j{}^i + \left(K^h_h\right)^2\right)$$

$$\alpha\, S^{\circ}_{i} \equiv \overline{\nabla}_h\, K_i{}^h - \partial_i\, K^h_h$$

Using (3-1) the other terms in A°_{o} and $\alpha\, A^{\circ}_{i}$ can also be expressed in terms only of the geometric elements $\overline{g}_t$ and K_t on S_t.

The intrinsic Cauchy data on a slice S_o for gravitation with Gauss-Bonnet terms are like for usual gravity a metric and a symmetric 2-tensor on S_o, satisfying the constraints,

$$A_{\perp\perp} \equiv A_{\alpha\beta}\, n^{\alpha}\, n^{\beta} = 0 \qquad \text{"hamiltonian" constraint}$$

$$A^{\parallel}_{\perp} = (A_{\perp i}) = \pi^{\beta}_{i}\, n_{\alpha}\, A^{\alpha}{}_{\beta} = -\alpha\, A^{\circ}_{i} = 0 \qquad \text{"momentum" constraint.}$$

An analytic solution of the equations $A_{ij} = 0$ on $U = S \times I$ which satisfies the constraints on S_o satisfies the equations $A_{\alpha\beta} = 0$ in a neighborhood of S_o, for any analytic choice of lapse (and shift if we introduce it) due to the identities (1-3) which are then a first order homogeneous system of the Cauchy Kovalevski type for the d quantities A°_{α}.

The same is true if, instead of $A_{ij} = 0$, we consider the equations

58

$$\tilde{A}_{ij} \equiv A_{ij} - \frac{1}{d-2} g_{ij} A^{\alpha}_{\alpha} \equiv R_{ij} - \frac{2}{d-2} \Lambda g_{ij} + B_{ij} - \frac{1}{d-2} g_{ij} B^{\alpha}_{\alpha} = 0 \quad .$$

IV - Evolution. Analytic case.

The equations $A_{ij} = 0$, or $\tilde{A}_{ij} = 0$ are, when the lapse is given as well as the shift (here taken to be zero) a system of non linear $d(d-1)/2$ partial differential equations for the unknown g_{hk}. It results from §3, as already remarked by Aragone that they are linear in the second derivatives $\partial^2_{oo} g_{hk}$. They are of the Cauchy Kovalevski type in a neighborhood in $S \times \mathbb{R}$ of the manifold $S_o = S \times \{0\}$, for the Cauchy data $\bar{g}$ and K, if they can be solved with respect to the $\partial^2_{oo} g_{hk}$, that is if the determinant of the coefficients of these derivatives is non zero.

We denote by $\simeq$ equality modulo the addition of terms which contain no second derivatives $\partial^2_{oo} g_{hk}$. We have

$$R_{ij} \simeq \frac{1}{2\alpha^2} \partial^2_{oo} g_{ij}$$

$$B_{ij} \simeq \sum_{p=2}^{P} g_{j\ell} k_p (2p)^2 \mathcal{E}_i^{\ell \ell_2 \ell_3 \cdots \ell_{2p}}{}_{m_2 m_3 \cdots m_{2p}} R_{\ell_2 o}{}^{m_2 o} R_{\ell_3 \ell_4}{}^{m_3 m_4} \cdots R_{\ell_{2p-1} \ell_{2p}}{}^{m_{2p-1} m_{2p}}$$

$$\simeq \frac{1}{2\alpha^2} X^{hk}_{ij} \partial^2_{oo} g_{hk}$$

with

$$X^{hk}_{ij} = g_{j\ell} g^{mk} \sum_{p=2}^{P} k_p (2p)^2 \mathcal{E}_i^{\ell h \ell_3 \cdots \ell_{2p}}{}_{m m_3 \cdots m_{2p}} R_{\ell_3 \ell_4}{}^{m_3 m_4} \cdots R_{\ell_{2p-1} \ell_{2p}}{}^{m_{2p-1} m_{2p}} \quad .$$

The system $\tilde{A}_{ij} = 0$ is of the Cauchy-Kovalevski type, for the unknown g_{ij} and an arbitrarily given α if the determinant of the matrix M with elements (a capital index is a pair of ordered indices)

59

$$M_I^J \equiv M_{ij}^{hk} = \frac{1}{2\alpha^2}\left(\delta_I^J + \varkappa\, Y_{ij}^{hk}\right) \quad ,$$

$$Y_{ij}^{hk} \equiv X_{ij}^{hk} - \varkappa \frac{1}{d-2}\, g_{ij}\, g^{\ell m}\, X_{\ell m}^{hk} \quad ,$$

$$g^{\ell m}\, X_{\ell m}^{hk} \equiv g^{mk} \sum_{p=2}^{P} k_p (2p)^2 (d-1)\, \varepsilon_{m\, m_3 \cdots m_{2p}}^{h\, \ell_3 \cdots \ell_{2p}}\, R_{\ell_3 \ell_4}{}^{m_3 m_4} \cdots R_{\ell_{2p-1} \ell_{2p}}{}^{m_{2p-1} m_{2p}}$$

is non identically zero.

We have, $\mathbb{1}$ denoting the unit matrix and $D = d(d-1)/2$,

$$\det M = \left(\frac{1}{2\alpha^2}\right)^D \det (\mathbb{1} + \varkappa Y)$$

and

$$\det (\mathbb{1} + \varkappa Y) = 1 + \varkappa\, a_1 + \varkappa^2\, a_2 + \ldots + \varkappa^D\, a_D$$

with

$$a_1 = \delta_J^I\, Y_I^J = \operatorname{tr} Y = \delta_h^i\, \delta_k^i\, Y_{ij}^{hk} = X_{ij}^{ij} - \frac{1}{d-2}\, g_{ij}\, g^{\ell m}\, X_{\ell m}^{ij} =$$

$$= \sum_{p=2}^{P} k_p\, (2p)^2\, \varepsilon_{m_3 \cdots m_{2p}}^{\ell_3 \cdots \ell_{2p}}\, R_{\ell_3 \ell_4}{}^{m_3 m_4} \cdots R_{\ell_{2p-1} \ell_{2p}}{}^{m_{2p-1} m_{2p}}$$

$$a_2 = \varepsilon_{J_1 J_2}^{I_1 I_2}\, Y_{I_1}^{J_1}\, Y_{I_2}^{J_2} = (\operatorname{tr} Y)^2 - Y_{ij}^{hk}\, Y_{hk}^{ij}$$

$$\vdots$$

$$a_D = \varepsilon_{J_1 \cdots J_D}^{I_1 \cdots I_D}\, Y_{J_1}^{I_1} \ldots Y_{J_D}^{I_D} = \det Y$$

a_q is a polynomial in the components $R_{ij}{}^{k\ell}$ of the Riemann tensor of g which can be expressed on S_o, using (3-1), in terms of the Cauchy data $\bar{g}$ and K.

Theorem Let $(S, \bar{g}, K)$ be an analytic initial data set, satisfying the constraints and such that $\det(\mathbb{1} + \varkappa Y)_S \neq 0$. There exists an analytic space

time (V,g) taking these initial data and solution of the equations of stringy gravitation with Gauss-Bonnet terms.

The (analytic) lapse is arbitrary : to different choices of lapse correspond locally isometric space times.

Proof : The Cauchy-Kovalevski theorem.

Remark if χ, or if K and the Riemann tensor of $\bar{g}$, are small enough then $\det(\mathbb{1} + \chi\, Y)_S \neq 0$.

V - Characteristics as possible wave fronts.

In order to study possible propagation of Gauss-Bonnet gravity, and even tually get rid of the analyticity hypothesis in the solution of the Cauchy problem, we now look for the possible significant discontinuities of the second derivatives of the metric across a d-1 dimensional submanifold S of a given space time (V,g) solution of the equations of Gauss-Bonnet gravity. Such hypersurfaces are called wave fronts.

We know (cf Lichnerowicz[11]) that the significant discontinuities of the second derivatives of the metric - that is those which cannot be removed by a C^2 by pieces change of coordinates - are the discontinuities in $\partial^2_{oo}\, g_{ij}$ if S has local equation x^o = constant. The calculations made in the previous paragraph show that these discontinuities can occur across S if and only if

$$\det M|_S = 0 \quad . \tag{5-1}$$

In arbitrary coordinates, where the equation of S is $f(x^\alpha) = 0$, the condition for S to be a wave front reads

$$\left(- g^{\lambda\mu}\, \partial_\lambda f\, \partial_\mu f\right)^D \left(1 + \chi\, a_1 + \ldots + \chi^D\, a_D\right) \neq 0$$

where a_q is an invariant polynomial of order (P-1)q in the tensor ρ, projection on S of the Riemann tensor. It can be seen using the expression of

ρ and the antisymmetry of the $\mathcal{E}$ tensor that a_q is only a polynomial of order 2q in n. We find for instance when P = 2

$$a_1 = C\, S^{\alpha\beta}\, n_\alpha\, n_\beta \qquad , \qquad (\text{using } n^\alpha\, n_\alpha = -1)$$

(result found by Aragone[10])

and a_2 is of the form, with C, C_o, C_1, C_2 numbers depending on d

$$a_2 = C_o \left(\rho^\alpha_\alpha\right)^2 + C_1\, \rho^\lambda_\alpha\, \rho^\alpha_\lambda + C_2\, \rho_{\alpha\beta}{}^{\lambda\mu}\, \rho^{\alpha\beta}{}_{\lambda\mu}$$

which, using antisymmetries and $n^\alpha\, n_\alpha = -1$, reduces to a polynomial of degree 4 in n.

We obtain the equation for the wave fronts by replacing $n_\alpha\, n_\beta$ in a_q by

$$n_\alpha\, n_\beta = - \frac{\partial_\alpha f\, \partial_\beta f}{g^{\lambda\mu}\, \partial_\lambda f\, \partial_\mu f}$$

and we see that the hypersurface S, f = constant can be a wave front if

$$\Delta \equiv \left(- g^{\lambda\mu}\, \partial_\lambda f\, \partial_\mu f\right)^D + \chi \left(- g^{\lambda\mu}\, \partial_\lambda f\, \partial_\mu f\right)^{D-1} b_1(\nabla f)$$

$$+ \chi^2 \left(- g^{\lambda\mu}\, \partial_\lambda f\, \partial_\mu f\right)^{D-2} b_2(\nabla f) + \ldots + \chi^D\, b_D(\nabla f) = 0$$

where $b_q(\nabla f)$ is an homogeneous polynomial of degree 2q in ∇f, whose coefficients vanish when the curvature tensor vanishes.

The wave front cone at a point of V, is a cone in the cotangent space, obtained by replacing $\partial_\lambda f$ by a covariant vector ξ_λ, of degree 2D. By taking the parameter χ small - or the curvature small - it is possible to insure that this cone remains in a region close to the null cone of the metric, using the property that this null cone is convex (cf 12), however there is no reason to consider that the product of the null cone and the cone $g^{\lambda\mu}\, \xi_\lambda\, \xi_\mu - \chi\, b_1(\xi) = 0$ approximates the full cone C.

62

VI - Harmonic coordinates.

Coordinates are harmonic if the metric satisfies the conditions

$$F^\lambda \equiv g^{\alpha\beta}\,\Gamma_\alpha{}^\lambda{}_\beta = 0 \quad . \tag{6-1}$$

It is well known that

$$R_{\alpha\beta} \equiv -\frac{1}{2}\,g^{\lambda\mu}\,\partial_\lambda{}^2{}_\mu\,g_{\alpha\beta} + g_{\alpha\lambda}\,\partial_\beta F^\lambda + g_{\beta\lambda}\,\partial_\alpha F^\lambda + h_{\alpha\beta}(g,\partial g) \tag{6-2}$$

where $H_{\alpha\beta}$ depends only on the metric and its first derivative. We set

$$R_\alpha^{(h)}{}_\beta \equiv -\frac{1}{2}\,g^{\lambda\mu}\,\partial^2_{\lambda\mu}\,g_{\alpha\beta} + h_{\alpha\beta} \tag{6-3}$$

and

$$\tilde{A}_\alpha^{(h)}{}_\beta \equiv R_\alpha^{(h)}{}_\beta + \varkappa\,\tilde{B}_{\alpha\beta} \tag{6-4}$$

where we do not truncate $B_{\alpha\beta}$ by the use of (6-1).

We deduce from the conservation identities that a solution of $\tilde{A}_\alpha^{(h)}{}_\beta = 0$ satisfies the homogeneous wave equations in F^λ :

$$\nabla^\alpha\,\partial_\alpha F^\lambda = 0 \quad . \tag{6-5}$$

On the other hand a hyperbolic metric $g_{\alpha\beta}$ solution of $\tilde{A}_\alpha^{(h)}{}_\beta = 0$ which satisfies the constraints, on S_o ($x^\circ = 0$) satisfies also

$$\partial_o F^\lambda\big|_{x^\circ=0} = 0 \quad . \tag{6-6}$$

We deduce from this remark :

Proposition A solution of $A_\alpha^{(h)}{}_\beta = 0$ which satisfies the constraints on S_o, and

63

$$F^{\lambda}|_{S_o} = 0 \tag{6-7}$$

satisfies $F^{\lambda} = 0$ in all the future of S_o, determined by the isotropic cone of the metric, under only mild regularity hypothesis (as necessary for the uniqueness theorem for the wave equations), and hence satisfies $A_{\alpha\beta} = 0$.

In contradiction with $A_{\alpha\beta} = 0$, the system $\tilde{A}^{(h)}_{\alpha\ \beta} = 0$ is of the Cauchy-Kovalevski type. Its characteristic determinant is non zero except on a cone, the characteristic cone.

However this cone is not the light cone : the elements of the characteristic determinant are

$$P^{\rho\sigma}_{\alpha\beta} = \frac{\partial\, \tilde{A}^{(h)}_{\alpha\ \beta}}{\partial\left(\partial^2_{\lambda\mu}\, g_{\rho\sigma}\right)}\, \xi_\lambda\, \xi_\lambda = -\frac{1}{2}\, g^{\lambda\mu}\, \xi_\lambda\, \xi_\mu\, \delta^\rho_\alpha\, \delta^\sigma_\beta \tag{6-8}$$

$$+ \chi\, \frac{\partial\, B_{\alpha\beta}}{\partial\left(\partial^2_{\lambda\mu}\, g_{\rho\sigma}\right)}\, \xi_\lambda\, \xi_\mu$$

where rows and columns are numbered by ordered pairs of indices $(\alpha\beta)$, $(\rho\sigma)$.

Proposition (cf 12) The characteristic determinant of the system

$$\tilde{A}^{(h)}_{\alpha\ \beta} = 0 \tag{6-9}$$

is, with $C = d^{-d(d+1)/2}$

$$\det P^{(h)} = C\, \left(g^{\lambda\mu}\, \xi_\lambda\, \xi_\mu\right)^d\, \Delta$$

where Δ is the polynomial giving the wave front cone determined in the previous paragraph.

Writing the Gauss-Bonnet gravity in harmonic coordinates as $A^{(h)}_{\alpha\beta} = 0$ introduces the isotropic cone as a spurious wave front cone, but preserves the true one.

64

Without further information on this cone : reality, simplicity, convexity, it is not easy to give more results on the general Cauchy problem for the classical system of partial differential equations of Gauss-Bonnet gravity.

VII - Shocks and High frequency waves.

The equations of gravitation with Gauss-Bonnet terms are fully non linear ; as such they offer a new challenge to the specialists in partial differential equations, in particular the study of shock waves as well as the determination of high frequency waves by asymptotic expansion could be untractable . However the non linear terms have a remarquable property , a consequence of which is the result obtained by Boulware and Deser that gravitational plane waves are also solution of the equations with Gauss-Bonnet terms . We now give this remarkable property in its full generality.

Theorem 1. If the Riemann tensor is of the pure radiative form , then the Gauss Bonnet correction in the equations $A_{\alpha\beta} = 0$ is identically zero .

Proof : The Riemann tensor is called purely radiative if it reads (Lichnerowicz 1961)

$$R_{\alpha\beta,\lambda\mu} = h_{\alpha\lambda}\, n_\beta n_\mu - h_{\beta\lambda} n_\alpha n_\mu + h_{\beta\mu}\, n_\alpha n_\mu - h_{\alpha\mu} n_\beta n_\lambda$$

The result follows from the antisymmetries of the Kronecker tensor .

The theorem 1 shows that if we try to determine shocks as solutions which admit discontinuities in the first derivatives of the metric across a hypersurface S we shall indeed find equations which do not contain the meaningless square of the measure $\delta(S)$. Unfortunately these equations still appear to contain in general the product of $\delta(S)$ by a function which is discontinuous across S . This is not defined , but it has been shown in special cases that such products in fact do not occur[21] and the study of the generic case remains to be done . We shall not pursue this way here ,

but instead we shall give some results about high frequency waves : it is a more flexible subject and perhaps more interesting physically in this context .

A metric g is said to represent an high frequency wave if it depends on the point x of the manifold V with two different scales : it is defined through a mapping from V x R into the space of metrics on V ,

$$(x , \xi) \mapsto g(x , \xi) , x \in V , \xi \in R$$

and by replacing ξ by the product $\omega \varphi(x)$, where ω is a large parameter and φ is a scalar function on V . With this hypothesis the partial derivatives of the components of g in local coordinates are given by , with obvious notations ,

$$\partial_\lambda g_{\alpha\beta} = \underline{\partial}_\lambda g_{\alpha\beta} + \omega\, g'_{\alpha\beta} n_\lambda \quad , \quad n_\lambda = \partial_\lambda \varphi \quad , \quad g'_{\alpha\beta} = \partial g_{\alpha\beta} / \partial \xi$$

and make appear , by choice of a large enough ω , a rapid variation of g in the direction n transversal to the submanifolds φ = constant , called submanifolds of constant phase , or sometimes wave fronts .

We then look for a solution of the system of partial differential equations under the form of a formal series in inverse powers of ω . When we report formally this series in the equations we obtain another formal series in these inverse powers and we say that we have an asymptotic solution of order p if all coefficients of powers of ω vanish up to the coefficient of ω^{-p} . An asymptotic solution is also an approximate solution if the remainder satisfies appropriate boundedness conditions .

There are several possible choices , already in general relativity, to look for these asymptotic expansions starting from a metric independant of ω called a background , and perturbating it either with terms of order ω^{-2} (cf 13, 16) or by terms of order ω^{-1} which induce a "back reaction" (cf 14, 15) : in this case the background metric cannot be a solution of the vacuum Einstein equations. In the case of the equations with Gauss Bonnet terms the situation is enriched by the presence of the coupling constant . We shall present below some results which take advantage of this possibility .

66

We consider the case where the coupling constant between the usual Einstein tensor and the Gauss Bonnet terms is of order ω^{-1} , for simplicity in writing we take only the quadratic polynomial in the Riemann curvature : the results are essentially the same if higher order terms are also considered . The equations read :

$$A_{\alpha\beta} \equiv S_{\alpha\beta} + \omega^{-1} B_{\alpha\beta} \quad , \quad B_{\alpha\beta} \equiv \epsilon^{\gamma\lambda_1\lambda_2\lambda_3\ \lambda_4}_{\alpha\mu_1\mu_2\mu_3\mu_4} R_{\lambda_1\lambda_2}{}^{\mu_1\mu_2} R_{\lambda_3\lambda_4}{}^{\mu_3\mu_4} \quad .$$

We look for a solution which is a perturbation of order ω^{-2} of some given metric $\underline{g}$ called background , independant of ω

$$g_{\alpha\beta}(x\ ,\ \omega\ \varphi(x)) \equiv \underline{g}_{\alpha\beta}(x) + \omega^{-2} h_{\alpha\beta}(x\ ,\ \omega\ \varphi(x)) + \omega^{-3} \kappa_{\alpha\beta}(x\ ,\ \omega\ \varphi(x))$$

We suppose that h and k are uniformly bounded independently of ω as well as their first and second derivatives with respect to the variables x and $\xi = \omega\varphi$, so that this expansion is meaningfull, as well as the ones we shall now compute.

We first deduce from the above formula

$$g^{\alpha\beta} \equiv \underline{g}^{\alpha\beta} - \omega^{-2} h^{\alpha\beta} + \omega^{-3} K^{\alpha\beta} \ , \ h^{\alpha\beta} \equiv \underline{g}^{\alpha\rho}\ \underline{g}^{\beta\sigma} h_{\rho\sigma} \ , \ \kappa^{\alpha\beta} \text{ bounded} \quad .$$

We find for the Christoffel symbols

$$\Gamma^{\lambda}_{\alpha\beta} \equiv \underline{\Gamma}^{\lambda}_{\alpha\beta} + \omega^{-1}\ \gamma^{\lambda}_{\alpha\beta} + \omega^{-2}\ N^{\lambda}_{\alpha\beta}$$

where

$$\gamma^{\lambda}_{\alpha\beta} \equiv \tfrac{1}{2}\ \underline{g}^{\lambda\mu}\ (\ h'_{\alpha\mu} n_\beta + h'_{\beta\mu} n_\alpha - h'_{\alpha\beta} n_\mu\ )$$

which gives for the Riemann tensor

$$R_{\alpha\beta}{}^{\lambda}{}_{\mu} \equiv \underline{R}_{\alpha\beta}{}^{\lambda}{}_{\mu} + r_{\alpha\beta}{}^{\lambda}{}_{\mu} + \omega^{-1}\ T_{\alpha\beta}{}^{\lambda}{}_{\mu} + \omega^{-2}\ M_{\alpha\beta}{}^{\lambda}{}_{\mu} \quad ,$$

where r is of the purely radiative form defined previously : its components beeing raised or lowered with the background metric we have

$$r_{\alpha\beta\,.\,\lambda\mu} = \tfrac{1}{2}\,(\,h''_{\beta\lambda}n_\mu - h''_{\beta\mu}n_\lambda)\,n_\alpha - (h''_{\lambda\alpha}n_\mu - h''_{\alpha\mu}n_\lambda)\,n_\beta \quad .$$

The Ricci tensor is then given by

$$R_{\alpha\beta} \equiv \underline{R}_{\alpha\beta} + r_{\alpha\beta} + \omega^{-1}\,T_{\alpha\beta} + \omega^{-2}\,N_{\alpha\beta}$$

where underlines quantities are relative to the background metric $\underline{g}$ and $r_{\alpha\beta}$ is given by

$$r_{\alpha\beta} \equiv \tfrac{1}{2}\,\underline{g}^{\lambda\mu}\,(\,h''_{\lambda\beta}\,n_\alpha\,n_\mu + h''_{\lambda\alpha}n_\beta\,n_\mu - h''_{\lambda\mu}\,n_\alpha\,n_\beta - h''_{\alpha\beta}\,n_\lambda\,n_\mu\,)$$

We deduce from these expressions that the Gauss Bonnet equations admit the expansion , whose coefficients of powers of ω are uniformly bounded,

$$A_{\alpha\beta} \equiv \underline{S}_{\alpha\beta} + r_{\alpha\beta} - \tfrac{1}{2}\underline{g}_{\alpha\beta}r + \omega^{-1}\,(T_{\alpha\beta} - \tfrac{1}{2}\underline{g}_{\alpha\beta}T + 2\underline{g}_{\beta\gamma}\underline{\epsilon}^{\gamma\lambda_1\lambda_2\lambda_3\lambda_4}_{\alpha\mu_1\mu_2\mu_3\mu_4}\underline{R}_{\lambda_1\lambda_2}{}^{\mu_1\mu_2}\,r_{\lambda_1\lambda_2}{}^{\mu_1\mu_2}) + \omega^{-2}M_{\alpha\beta}$$

To obtain this formula we have used the theorem 1 which shows that the quadratic term in the radiative part r vanishes identically. We see that if the high frequency wave is to satisfy the equations $A_{\alpha\beta} = 0$ at the first order of approximation in ω^{-1} then we must have

$$\underline{S}_{\alpha\beta} + r_{\alpha\beta} - \tfrac{1}{2}\,\underline{g}_{\alpha\beta}r = 0 \quad , \text{ or equivalently, } \quad \underline{R}_{\alpha\beta} + r_{\alpha\beta} = 0$$

These equations are identical with the equations obtained in ordinary Einstein gravity when looking for high frequency waves at zero order in ω from known results in this case (cf 18) we deduce the following

Theorem 2. The high frequency wave is a significant perturbation of the background metric, solution of order zero in ω of the Einstein equations with a Gauss-Bonnet perturbation of order -1 in ω if and only if

68

1. The background metric is a solution of the vacuum Einstein equations $\underline{R}^{\alpha\beta} = 0$, and the wave fronts are null hypersurfaces of this metric,

2. The perturbation $h_{\alpha\beta}$ satisfies the relations

$$n_\alpha\, h - 2\, n^\lambda\, h_{\alpha\lambda} = 0 \quad , \quad \text{with } h = g^{\lambda\mu}\, h_{\lambda\mu} \quad , \quad n^\lambda = g^{\lambda\mu}\, n_\mu \quad .$$

Note : a perturbation is called significant if it cannot be made to vanish by a high frequency change of coordinates of the same type as the wave itself.

Radiative coordinates. We have denoted by "radiative" coordinates in which the wave fronts $\varphi(x)$ = constant are the hypersurfaces x^0 = constant. In such coordinates the above results read:

$$g^{00} = 0 \ , \ n_0 = 1 \ , \ n_i = 0 \ , \ n^0 = 0 \ , \ n^i = g^{0i} \text{ and}$$

$$n^j h_{ij} = 0 \quad , \quad g^{ij}\, h_{ij} = 0$$

We now look for the conditions which insure that our high frequency wave is a solution of the equations up to the order ω^{-2}. They will give us as usual in these problems propagation equations for the first order perturbation h.

The coefficients of the term in ω^{-1} in each equation is the sum of a term coming from the pure Einstein tensor and a term coming from the Gauss Bonnet correction. The first term has already been determined in previous works (cf 17, 18), in particular in radiative coordinates, it has been found that for the Ricci tensor these coefficients split between linear differential operators in the significant part of the perturbation h_{ij}

$$R^1_{ij} \equiv - n^\lambda \underline{\nabla}_\lambda\, h'_{ij} - \tfrac{1}{2}\, h'_{ij}\, \underline{\nabla}_\lambda n^\lambda$$

and terms $R^1_{0\alpha}$ which contain the non significant part $h_{0\alpha}$.

We shall now compute the Gauss Bonnet term. Due to the theorem 1 it will

69

be of the form

$$\epsilon \cdots \underline{R} \cdots \underline{R} \cdots + 2\epsilon \cdots \underline{R} \cdots r \cdots$$

The first term of the sum is independent of $\omega\varphi$, it must vanish if we want the perturbation to be uniformly bounded in ω. We compute the second term in radiative coordinates.

We deduce from the values of $\underline{g}$ and n in radiative coordinates that all components of $r_{\alpha\beta\lambda\mu}$ are zero, except those which contain two zero indices which are given by:

$$r_{i0j0} = \tfrac{1}{2}\, h''_{ij}$$

from this result we deduce that, for a perturbation satisfying the theorem 2, all the components of $r_{\alpha\beta}{}^{\lambda\mu}$ are zero except those which contain only one zero index, situated moreover in a bottom position, and then given by :

$$r_{0i}{}^{hk} = \tfrac{1}{2}\,\left(n^h\, \underline{g}^{kj} - n^k\, \underline{g}^{hj} \right) h''_{ij}$$

We deduce from these facts that if we set

$$\Theta^{\alpha}_{\gamma} \equiv 2\,\epsilon^{\gamma\lambda_1\ldots\lambda_4}_{\alpha\mu_1\ldots\mu_4}\, \underline{R}_{\lambda_1\lambda_2}{}^{\mu_1\mu_2}\; r_{\lambda_3\lambda_4}{}^{\mu_3\mu_4}$$

we have in radiative coordinates

$$\Theta^{\alpha}_{\gamma} = 4\,\epsilon^{\alpha 1_1 1_2 0\ i}_{\gamma\mu_1\mu_2 h\ k}\; r_{0i}{}^{hk}\; \underline{R}_{1_1\ 1_2}{}^{\mu_1\mu_2}$$

hence

$$\Theta^{0}_{\alpha} = 0 \quad , \quad \Theta^{j}_{0} = -\,8\,\epsilon^{j1pi}_{mqhk}\, n^h\, \underline{g}^{kn}\, h''_{in}\, \underline{R}_{1p}{}^{qh} \quad ,$$

$$\Theta^{c}_{j} = 16\,\epsilon^{c1pi}_{jmhk}\, n^h\, \underline{g}^{kn}\, h''_{in}\, \underline{R}_{1p}{}^{0m} \quad , \quad \Theta \equiv \Theta^{\alpha}_{\alpha} = \Theta^{j}_{j}$$

70

The equations satisfied by the significant part h_{ij} of the perturbation read

$$\overset{1}{R}_{ij} + g_{ic}\,\Theta^{c}_{j} - g_{ij}\,\frac{\Theta}{d-2} = 0 \qquad (P)$$

We shall enunciate as a theorem the results we have just obtained.

Theorem 3. For a general background metric g, a high frequency wave solution to order zero in ω of the Einstein equations with Gauss Bonnet terms (coupling constant $\chi = \omega^{-1}$) is also solution to order ω^{-1} if

1. The background metric annuls the Gauss Bonnet term.

2. The significant part of the perturbation h_{ij} satisfies the system of differential equations (P), called propagation equations. these are linear differential equations of the first order in h'_{ij} which contain both a derivation in the direction of the "light ray" n and a derivation with respect to the supplementary variable $\xi = \omega\varphi$.

The fact that the derivative h''_{ij} does appear in the propagation equation shows that signals will be, in the case of a generic background, distorted along their propagation (cf 19); the wave fronts $\varphi(x)$ = constant will not be exceptionnal in the sense of Lax and Boillat (cf 20).

Many developments can originate from this general study. But some idea of physical applications would be usefull both to motivate the effort and to guide the researcher. Physical situations where these partial differential equations could be relevant are considered by R. Kerner and collaborators (cf 8, 9, 22).

References

[1] M. Green, J. Schwarz, E. Witten, *String theory*, Cambridge University Press (1987).

[2] D. Lovelock, J. Maths Phys. 12 498 (1971).

[3] D. Boulware and S. Deser, Phys. Rev. D 30, 707 (1984).

[4] B. Zumino, in Physics Reports (1986).

[5] Y. Choquet-Bruhat and C. DeWitt, *Analysis manifolds and Physics*, Revised edition, North Holland 1982 p.394.
M. Dubois-Violette and J. Madore, Preprint Orsay 1986.

[6] N. Deruelle (private communication)

[7] D. Boulware and S. Deser, Phys. Lett. B, 55, 2656 (1985).

[8] R. Kerner, C.R. Ac. Sc. Paris, t.304, II, 621 (1987).

[9] B. Giorgini and R. Kerner, *Classical and quantum gravity* (to appear).

[10] C. Aragone, *The characteristics of stringy gravity*, Silog 5 Proceedings, M. Novello Ed., to appear, World Publishing C°.

[11] A. Lichnerowicz, *Théories Relativistes de la Gravitation et de l'Electromagnétisme*, Masson 1955.

[12] Y. Choquet-Bruhat, The Cauchy problem for stringy gravity, to appear 1988, J. of Math. Phys.

[13] A. Lichnerowicz, Ondes et radiations électromagnétiques et gravitationnelles en relativité générale. Annali di Matematica, 50, p.2-95, (1960).

[14] L. Garding, T. Kotake, J. Leray, Bull. Soc. Maths. 92 p.263-361, 1964.

[15] Y. Choquet-Bruhat, C.R. Ac. Sc. Paris, 258, p.3809-12, 1964.

[16] R. Isaacson, Phys. Rev. 166, 1263 and 1272, 1968.

[17] Y. Choquet-Bruhat, Comm. Maths. Phys., 12, p.16-35 (1969).

[18] Y. Choquet-Bruhat in "Ondes et Radiations Gravitationnelles" ed. C.N.R.S. Paris 1974 and in "Proceedings of the first Marcel Grossmann meeting" Trieste 1980, Denardo and Ruffini ed.

[19] Y. Choquet-Bruhat, J. Maths pures et appliquées, 48, p.117-158, 1969.

[20] G. Boillat, Ondes asymptotiques, Ann. Mat. pura ed Appl. 61, 1976, p.31-44.

72

[21] A. Tomimatsu, H. Ishihara, Preprint Nagoya University 1987.

[22] J. Fabris, R. Kerner, Generalization of supergravity in eleven dimension with the Gauss-Bonnet invariants. Helv. Phys. Acta (1988).

Institut de Mèchanique
Universitè Paris VI
4 place Jussieu
75230 PARIS CEDEX 05
FRANCE

RENDICONTI DEL CIRCOLO MATEMATICO DI PALERMO
Serie II, Suppl. 45 (1996), pp. 111-123

INTERACTION OF GRAVITATIONAL AND FLUID WAVES.

Yvonne Choquet-Bruhat and Antonio Greco

Introduction.

We will study the interaction of gravitational waves with perfect fluids using the formalism of high frequency asymptotic expansions initiated by Keller and Lax[1], developped in the linear case by Garding, Kotake and Leray[2] and in the nonlinear case by C-B[3]. This formalism has already had many applications to the construction of fast oscillating solutions of partial differential equations of physical interest. In particular it has been applied to relativistic perfect fluids [3] and to the coupled Dirac and Yang-Mills equations [7]. The study of high frequency gravitational waves, and their coupling with electromagnetic[4,5] or scalar[6] waves has also been made. However in the case of Einstein equations there is a problem of choice of gauge: a high frequency perturbation is not covariantly defined. We will here study the coupling of high frequency fluid and gravitational waves, using gauge conditions inspired by the hyperbolic formulation of the Einstein equations given by C-B-Ruggeri[8] and C.B-York [9], where the physical gravitational unknown, covariant by space diffeomorphisms , are the metric and extrinsic curvature of the space slices.

Definitions.

J. Leray, generalizing P. Lax previous definitions, has introduced for linear partial differential equations on $\mathbb{R}^n$ with vector valued unknown the definition of an asymptotic high frequency wave as a formal series of negative powers of a large parameter ω, analogous to a frequency, with coefficients depending on the one hand directly on the point x of $\mathbb{R}^n$ on the other hand on the product $\omega\varphi(x)$ where φ is a real valued scalar function called the phase. These coefficients must be bounded in some sense for the expansion to be meaningful. This formal series is an asymptotic solution if the formal series in ω obtained from the action of the partial differential operator on the asymptotic wave is identically zero. By truncating the asymptotic series for the unknown one may obtain approximate solutions of arbitrary order in ω.

In the case of a non linear P.D.E. it is not in general possible to obtain a meaningful asymptotic solution, but one can construct high frequency perturbations which are approximate solutions in a precise sense. Let D be a non linear partial differential operator acting on an unknown f, for instance a set of tensors on a manifold V. We say that the finite sum

$$f(x) = \underline{f}(x) + \omega^{-p} f^{(p)}(x,\xi)_{\xi=\omega\varphi(x)} + \omega^{-p-1} f^{(p+1)}(x,\xi)_{\xi=\omega\varphi(x)}$$

is a high frequency wave of order p if $\underline{f}$ is a set of tensors on V independent of ω and bounded in some norm, while $f^{(p)}$ and $f^{(p+1)}$ are sets of tensors on V× R bounded in the norm on V, uniformly on R (i.e. for all values of ξ). Such a high frequency wave is said to be an asymptotic solution of order q of the non linear differential system $D(f) = 0$ if it satisfies an equation of the form

$$D(f)(x) = \omega^{-q} F(x,\xi)_{\xi=\omega\varphi(x)}$$

where F is bounded in a norm on V, uniformly in ξ.

Remarks

1. The partial derivatives of a function $f(x,\omega\varphi(x))$ are given by the formula

$$\frac{\partial f}{\partial x^\alpha} = \partial_\alpha f + \omega f' \varphi_\alpha$$

where ∂_α is the partial derivative with respect to x^α for fixed ξ and

$$f' = \frac{\partial f}{\partial \xi}, \qquad \varphi_\alpha = \partial_\alpha \varphi$$

2. The uniform boundedness of f in ξ requires that f' satisfies an inequality of the type

$$\lim_{T\to\infty} \frac{1}{T} \int_0^\infty f'(.,\xi) d\xi = 0.$$

Einstein Equations with Source a Perfect Fluid.

We consider a pseudo riemannian metric of lorentzian signature on a 4 manifold $M \times \mathbb{R}$ written in a frame with time axis orthogonal to the space slices $M \times \{t\}$

$$-N^2(\theta^0)^2 + g_{ij}\theta^i\theta^j$$

The coframe θ^{α} is:

$$\theta^0 = dt, \quad \theta^i = dx^i + \beta^i dt$$

N and β are respectively a scalar function and a tangent vector on each M_t, i.e. t dependent scalar function and vector on M; they are called the lapse and the shift; $\bar{g}$ with components g_{ij} is a t dependent properly riemannian metric on M:

$$\bar{g} = g_{ij}dx^i dx^j$$

We introduce the differential operator $\hat{\partial}_0$ mapping t dependent tensors on M into t dependent tensors of the same type given by

$$\hat{\partial}_0 \equiv \partial_t - \mathcal{L}_\beta, \quad \mathcal{L}_\beta \text{ the Lie derivative with respect to } \beta$$

We denote by K the second fundamental form of M_t as a submanifold of the space time, it is a t dependent 2 tensor on M. We have

$$\hat{\partial}_0 g_{ij} = -2NK_{ij}. \tag{1}$$

We denote by H the trace of K:

$$H = g^{ij}K_{ij}.$$

We denote by an overbar the covariant derivatives and the Ricci tensor relative to the (t dependent) space metric $\bar{g}$.
The Einstein equations

$$R_{\alpha\beta} = \rho_{\alpha\beta}$$

together with known identities give the following equations:

$$R_{00} = N(\bar{\nabla}^i\partial_i N - NK_{ij}K^{ij} + \hat{\partial}_0 H) = \rho_{00} \tag{2}$$

$$R_{0j} = N(-\bar{\nabla}_h K_j^h + \partial_j H) = \rho_{0j} \tag{3}$$

$$R_{ij} = -N^{-1}\hat{\partial}_0 K_{ij} + HK_{ij} - 2K_{im}K_j^m - N^{-1}\bar{\nabla}_j\partial_i N + \bar{R}_{ij} = \rho_{ij}. \tag{4}$$

We take for the source ρ a perfect fluid, i.e.

$$\rho_{\alpha\beta} \equiv (e+p)u_\alpha u_\beta + \tfrac{1}{2}(e-p)g_{\alpha\beta}$$

with an equation of state linking the specific internal energy e with the pressure p and the entropy S:

$$e = f(p,S)$$

The motion of the fluid is then governed by the equations

$$M^\beta \equiv (e+p)u^\alpha \nabla_\alpha u^\beta + (g^{\alpha\beta} + u^\alpha u^\beta)\partial_\alpha p = 0$$

(6) $$M^4 \equiv \frac{\partial f}{\partial p}\, u^\alpha \partial_\alpha p + (e+p)\nabla_\alpha u^\alpha = 0$$

(7) $$M^5 \equiv u^\alpha \partial_\alpha S = 0.$$

High Frequency Asymptotic Expansions.

We consider the lapse N and shift β as gauge variables: they are a given scalar and vector field on M, dependent on t. The physical gravitational variables are the t dependent 2 tensors $\bar{g}$ and K. We suppose that $\bar{g}$, u, S and p admit high frequency asymptotic expansion around underlined backgound values given by

$$g_{ij}(x, \omega\varphi(x)) = \underline{g}_{ij}(x) + \{\omega^{-1} h_{ij}(x,\xi) + \omega^{-2} l_{ij}(x,\xi)\}_{\xi=\omega\varphi(x)}.$$

Then g^{ij} admits the expansion, where indices are lifted with $\underline{g}^{ij}$ as in all that follows,

$$g^{ij} = \underline{g}^{ij} - \omega^{-1} h^{ij} + \cdots$$

It results then from (1) that K admits the expansion

$$K_{ij} = \underline{K}_{ij}(x) + \{k_{ij}(x,\xi) + \omega^{-1}\chi_{ij}(x,\xi)\}_{\xi=\omega\varphi(x)}$$

with

$(1)_0$ $$\varphi_0 h'_{ij} = -2Nk_{ij}.$$

The connection coefficients of the space time metric in the chosen given frame are equal to their background value except

$$\omega^0_{ij} = -\,N^{-1}K_{ij}, \qquad \omega^i_{h0} = -Ng^{ij}K_{jh} = \omega^i_{0h} - \partial_{\ h}\beta^i, \quad \omega^i_{jh} = \Gamma^i_{jh}$$

their asymptotic expansions result from these formulas. In particular:

$$\omega^i_{jh} = \underline{\Gamma}^i_{jh} + \{\gamma^i_{jh}(x,\xi) + \omega^{-1}\gamma^i_{jh}{}^{(1)}(x,\xi) + \cdots\}_{\xi=\omega\varphi(x)} + \cdots$$

with, setting $f_{(ij)} = f_{ij} + f_{ji}$, and using $(1)_0$

$$\gamma^i_{jh} = -N\varphi_0^{-1}(\varphi_{(h}k^i_{j)} - \varphi^i k_{jh}).$$

We set also

$$u^\alpha(x,\omega\varphi(x)) = \underline{u}^\alpha(x) + \{\omega^{-1}v^\alpha(x,\xi) + \omega^{-2}w^\alpha(x,\xi)\}_{\xi=\omega\varphi(x)}$$
$$p(x,\omega\varphi(x)) = \underline{p}(x) + \{\omega^{-1}\pi(x,\xi) + \omega^{-2}\zeta(x,\xi)\}_{\xi=\omega\varphi(x)}$$
$$S(x,\omega\varphi(x)) = \underline{S}(x) + \{\omega^{-1}\sigma(x,\xi) + \omega^{-2}\Sigma(x,\xi)\}_{\xi=\omega\varphi(x)}.$$

Scalars, vectors and tensors in these formulas are supposed to be uniformly bounded in ξ.

Algebraic Conditions on the Gravitational Wave.

We consider the equation (2). Its highest order term is in ω. Its vanishing implies that, if $\varphi_0 \neq 0$,

$$(2)_{-1} \qquad \underline{g}^{ij}k'_{ij} = 0$$

The vanishing of the highest order term, in ω^1, of the equation (3) gives using $(2)_{-1}$

$$(3)_{-1} \qquad \varphi^i k'_{ij} = 0.$$

We now consider (4) at its highest order, ω. We have

$$\bar{R}^{(-1)}_{ij} \equiv \varphi_h \gamma'^h_{ij} - \varphi_i \gamma'^h_{hj}$$

A straightforward computation using $(2)_{-1}, (3)_{-1}$ gives :

$$\bar{R}^{(-1)}_{ij} = \varphi_0^{-1} N \varphi^h \varphi_h k'_{ij} \ .$$

The equations (4) imply therefore

$$R^{(-1)}_{ij} \equiv \mathrm{N}\{-N^{-2}(\varphi_0)^2 + \underline{g}^{ij}\varphi_i\varphi_j\}k'^{\cdot}_{ij} = 0$$

These equations imply either

(i) φ is a solution of the eikonal equation of the background: $\varphi^\alpha\varphi_\alpha = 0$, or

(ii) $k_{ij} = 0$.

We will say when $k_{ij} \neq 0$ that the gravitational wave is dominant: though it is of order 1 like the fluid wave, its "energy" is in fact of order zero in ω, like k, while the energy of the fluid wave is of order ω^{-1}.

Remark. Our gauge choice, N and β non oscillating, eliminates the perturbations of the form $h_{\alpha\beta}$ which can be made to vanish by an oscillating change of variables[4]. Such perturbations are shown by Ehlers and al.[13] to be the only possible ones with non isotropic phase for the linearized Einstein equations coupled with pressureless matter (dust).

Dominant Gravitational Wave. Algebraic Conditions on the Fluid Variables

The vanishing at the highest order, ω^0, of the fluid equations implies on the one hand that $\underline{u}^\alpha$, $\underline{p}$, $\underline{S}$ satisfy the background fluid equations, on the other hand it implies a linear system in $(v', \pi', \sigma') = (v'^I)$, I = 0, 1, ...,5

$$D^J_I v'^I = A^J$$

which reads

$$(5)_0 \qquad (\underline{e}+\underline{p})\underline{u}^\alpha\varphi_\alpha v'^\beta + (\underline{g}^{\alpha\beta} + \underline{u}^\alpha\underline{u}^\beta)\varphi_\alpha\pi' = -(\underline{e}+\underline{p})\underline{u}^\alpha\underline{u}^\lambda\gamma^\beta_{\alpha\lambda} \equiv A^\beta$$

$$(6)_0 \qquad \underline{f}'_{\underline{p}}\underline{u}^\alpha\varphi_\alpha\pi' + (\underline{e}+\underline{p})\varphi_\alpha v'^\alpha = -(\underline{e}+\underline{p})\underline{u}^\lambda\gamma^\alpha_{\alpha\lambda} \equiv A^4$$

$$(7)_0 \qquad u^\alpha\varphi_\alpha\sigma' = 0.$$

A straightforward computation, using the values of the γ and the condition $(2)_{-1}$ gives

$$A^4 = 0, \quad A^0 = N^{-1}(\underline{e}+\underline{p})\underline{u}^i\underline{u}^j k_{ij}, \quad A^h = -N^2\varphi_0^{-1}\varphi^h A^0 + 2N(\underline{e}+\underline{p})\underline{u}^o\underline{u}^i k^h_i.$$

The determinant of the homogeneous operator on the left hand side is

$$D \equiv -(\underline{e}+\underline{p})^4(\underline{u}^\alpha\varphi_\alpha)^4(\underline{g}^{\alpha\beta} + (1-\underline{f}'_{\underline{p}})\underline{u}^\alpha\underline{u}^\beta)\varphi_\alpha\varphi_\beta.$$

Different cases occur, depending if this determinant vanishes or not for the considered phase φ.

We have seen that the hypothesis $k_{ij} \not\equiv 0$, case of a dominant gravitational wave implies that the submanifolds of constant "phase", $\varphi = constant$, must be isotropic, i.e.

$$\varphi^\alpha \varphi_\alpha \equiv \underline{g}^{\alpha\beta} \varphi_\alpha \varphi_\beta = 0,$$

the term k_{ij} being then left arbitrary by the highest order conditions, except for the restrictions $(2)_{-1}, (3)_{-1}$ found before.

The equation $(7)_0$ implies

$$\sigma' = 0.$$

Two cases occur for the restrictions on v and π.

(a) $\underline{f}'_p \neq 1$, compressible fluid.

(b) $\underline{f}'_p = 1$, incompressible, also called "hard" fluid.

Case (a). $\underline{f}'_p \neq 1$.

The determinant of D does not vanish. The system $(5)_0, (6)_0, (7)_0$ admits one and only one solution $(v', p', \sigma') = (V', P', \Sigma')$, depending linearly on its right hand side. This solution is given by

$$v'^\alpha = V'^\alpha \equiv \{\underline{u}^\beta \varphi_\beta (\underline{e} + \underline{p})\}^{-1} A^\alpha, \qquad p' = 0, \quad \sigma' = 0.$$

This solution is not identically zero if and only if

$$\underline{u}^i k_{ij} \not\equiv 0 \qquad \text{for some j}$$

Case (b). $\underline{f}'_p = 1$.

The determinant vanishes. The system has a solution if its right hand side satisfies

$$\ell_I A^I = 0 \quad \text{whenever } (\ell_I) = (\ell_\alpha, \ell_4, 0) \text{ satisfies } \ell_J D^J_I = 0.$$

We find here, up to a common factor,

$$\ell_\alpha = \varphi_\alpha, \quad \ell_4 = -a, \text{ with } \quad a \equiv \underline{u}^\alpha \varphi_\alpha$$

and we see that, if φ satisfies the eikonal of the background metric, while k satisfies the condition $(3)_{-1}$, then the solvability condition is satisfied.

The general solution at this order is therefore

$$v'^\alpha = d^\alpha U' + V'^\alpha, \qquad \pi' = d^4 U' + P', \quad \sigma' = 0.$$

U is an arbitrary function of x and ξ (except for boundedness conditions to insure the asymptotic validity), d^I is the right eigen vector of D depending only on x. It is found to be, up to a common factor

$$d^\alpha = a\underline{u}^\alpha + \varphi^\alpha, \qquad d^4 = -(\underline{e} + \underline{p})a, \qquad d^5 = 0.$$

while $(V'^\alpha, P', 0)$is a particular solution of the system, for which we can take P' = 0 and V' given by the same formula as in (a).
We have proved:

Theorem: *In a perfect fluid a high frequency gravitational wave of order one generates if* $u^i k_{ij} \not\equiv 0$ *for some j a fluid wave of order one and of null phase, valid at order zero.*

Propagation of the Dominant Gravitational Wave.

The vanishing of the term of order ω^{-1} in (1) is equivalent to

$$(1)_1 \qquad \hat{\partial}_0 h_{ij} + \varphi_0 l'_{ij} = -2N\chi_{ij}$$

while, using $(2)_{-1}$

$$H = \underline{H} + \omega^{-1}\{\underline{g}^{ij}\chi_{ij} - h^{ij}(\underline{K}_{ij} + k_{ij}\} + \omega^{-2}H^{(2)}.$$

The vanishing of the term of order zero in (2) implies then, using $(1)_0$, $(2)_0$, $(2)_{-1}$ the equation

$$\underline{R}_{00} - \underline{\rho}_{00} = N^2 k.k - N\varphi_0\{\underline{g}.\chi' - (h.k)'\}.$$

The asymptotic validity implies on the one hand that the background metric must satisfy the equation

$$\underline{R}_{00} - \underline{\rho}_{00} = N^2 \lim_{T\to\infty} \frac{1}{T}\int_0^T k(x,\xi).k(x,\xi)d\xi \equiv (\varphi_0)^2\mathcal{E}.$$

This expression gives a contribution of the gravitational wave to the zero order stress energy tensor. It expresses in our context a "back reaction" of the wave on the background.
The asymptotic validity implies on the other hand the linear equation in χ', with uniformly bounded solution

$$(2)_0 \qquad \underline{g}.\chi' = (h.k)' + \nu k.k - \nu^{-1}\mathcal{E}$$

where

$$\nu \equiv N\varphi_0^{-1}.$$

The validity at order zero of (3) gives, using $(2)_{-1}$

$$\underline{R}_{i0} = \underline{\rho}_{i0} + \varphi_i\varphi_0\mathcal{E}$$

and

$$\varphi^i\chi'_{ij} - \varphi_j\underline{g}.\chi' = -\underline{\nabla}_h k_j^h + \varphi_j\{-(h.k)' + \nu k.K - \nu k.k) - \nu^{-1}\mathcal{E})\}.$$

Using $(2)_0$ this equation reads

$(3)_0$ $$\varphi^i\chi'_{ij} = \varphi_j\nu k.\underline{K} - \underline{\nabla}_i k_j^i.$$

The validity at order zero of (4) is equivalent to

$$R_{ij}^{(0)} \equiv \underline{R}_{ij} + \bar{R}_{ij}^{(0)} - \bar{\underline{R}}_{ij} - N^{-1}(\varphi_0\chi'_{ij} + \hat{\partial}_0 k_{ij}) - 2k_{im}k_j^m - 2\underline{K}_{(i}^m k_{j)m}$$
$$+ \quad N^{-1}\gamma_{ij}^m\partial_m N = \underline{\rho}_{ij}.$$

We now use the formula

$$\bar{R}_{ij}^{(0)} \equiv \varphi_h\gamma'^{(1)h}_{ij} - \varphi_i\gamma'^{(1)h}_{jh} + \bar{\underline{\nabla}}_h\gamma_{ij}^h - \bar{\underline{\nabla}}_i\gamma_{jh}^h + \gamma_{ij}^h\gamma_{hm}^m - \gamma_{ih}^m\gamma_{jm}^h + \bar{\underline{R}}_{ij}$$

where $\gamma^{(1)}$, term of order 1 in the space connection, is given by

$$\gamma_{ij}^{(1)h} = \tfrac{1}{2}\{\varphi_{(i}l'^h_{j)} - \varphi^h l'_{ij} + \bar{\underline{\nabla}}_{(i}h_{j)}^h - \bar{\underline{\nabla}}^h h_{ij} - h^{hm}(\varphi_{(i}h'_{j)m} - \varphi_m h'_{ij})\}$$

and we use $(1)_0$ to obtain the formula

$$\gamma'^{(1)h}_{ij} = +\varphi_0^{-1}\{\varphi_{(i}\underline{g}^{hm}\hat{\partial}_0(\nu k_{j)m}) - \varphi^h\hat{\partial}_0(\nu k_{ij})\} - \bar{\underline{\nabla}}_{(i}(\nu k_{j)}^h) + \bar{\underline{\nabla}}^h(\nu k_{ij})$$
$$+\nu\{h^{hm}(\varphi_{(i}k_{j)m} - \varphi_m k_{ij})\}' - \{\varphi_{(i}\nu\chi'^h_{j)} - \varphi^h\nu\chi'_{ij}\}.$$

Using the conditions $(2)_{-1}$ and $(3)_{-1}$ these formulas imply

$$\varphi_h\gamma'^{(1)h}_{ij} = \nu\varphi^h\varphi_h\chi'_{ij} + \nu\varphi_{(i}\bar{\underline{\nabla}}_h k_{j)}^h + \varphi_h\bar{\underline{\nabla}}^h(\nu k_{ij}) - \varphi_0^{-1}\varphi^h\varphi_h\hat{\partial}_0(\nu k_{ij})$$
$$-2\nu^2\underline{K}.k\ \varphi_i\varphi_j - \nu\varphi_0^{-1}\varphi_{(i}k_{j)h}\hat{\partial}_0\varphi^h + \nu k_{(j}^h\bar{\underline{\nabla}}_{i)}\varphi_h$$

and

$$\gamma'^{(1)h}_{ih} \equiv -\varphi_i\nu\{\underline{g}.\chi' + 2\nu\underline{K}.k - (h.k)'\}$$

while

$$\gamma_{im}^h\gamma_{jh}^m = \nu^2(\varphi_i\varphi_j k.k - 2\varphi_h\varphi^h k_{im}k_j^m)\ .$$

Using these formulas $(2)_0$ and $(3)_0$ we find that the vanishing at order zero of (4) is equivalent to the equation

$$\underline{R}_{ij} - \underline{\rho}_{ij} - \varphi_i\varphi_j\mathcal{E} + (-N^{-1}\varphi_0 + \nu\varphi_h\varphi^h)\chi'_{ij} + (-\nu\varphi_0^{-1}\varphi_h\varphi^h - N^{-1})\hat{\partial}_0 k_{ij}$$

$$+2\nu\varphi^h\underline{\nabla}_h k_{ij} + 2k_{im}k_j^m(-1+\nu^2\varphi_h\varphi^h) + \{-\varphi_0^{-1}\varphi_h\varphi^h\partial_0\nu + \varphi^h\partial_h\nu + \underline{\nabla}_h(\nu\varphi^h)$$

$$+N^{-1}\nu\varphi^h\partial_h N\}k_{ij} - 2\underline{K}^h_{(i}.k_{j)h} - \nu\varphi_{(i}k_{j)h}(N^{-1}\partial^h N + \varphi_0^{-1}\hat{\partial}_0\varphi^h) = 0.$$

When the phase φ is isotropic, i.e. $\varphi^\alpha\varphi_\alpha \equiv \varphi^h\varphi_h - \nu^{-2} = 0$, the coefficients of χ'_{ij} and of the non linear term $k_{im}k_j^m$ vanish: the first result could have been foreseen on general grounds, the second is more remarkable: it expresses in our context that the Einstein equations satisfy a kind of "null condition", used in the proof of global existence of solutions of the Cauchy problem for data near to the Minkovskian initial data.
The coefficient of $\hat{\partial}_0 k_{ij}$ reduces to $2\nu\ \varphi^0$.
We see that the vanishing at order zero of (4) implies on the one hand

$$\underline{R}_{ij} = \underline{\rho}_{ij} + \mathcal{E}\varphi_i\varphi_j$$

and on the other hand a propagation equation for k of the type

$$2\{\varphi^0\hat{\partial}_0 k_{ij} + \varphi^h\bar{\underline{\nabla}}_h k_{ij}\} + Xk_{ij} + Y^h_{(i}k_{j)h} = 0,$$

i.e. an ordinary linear, homogeneous differential system of propagation along the ray φ^α.
From $(4)_0$, the expression of Y_i^h the definition of $\underline{K}$ in terms of $\hat{\partial}_0\underline{g}$ and the use of $(3)_1$ we deduce the equation

$$\varphi^\alpha\partial_\alpha(k.k) + Xk.k = 0$$

We had:

$$X = 2\nu^{-1}\varphi^\alpha\partial_\alpha\nu - \nu^{-1}\varphi^0\partial_0\nu + \bar{\underline{\nabla}}_h\varphi^h + N^{-1}\varphi^h\partial_h N$$

i.e. by a staightforward computation

$$X = 2\nu^{-1}\varphi^\alpha\partial_\alpha\nu + \underline{\nabla}_\alpha\varphi^\alpha.$$

The propagation equation for k.k reads therefore as the conservation law:

$$\underline{\nabla}_\alpha(\nu^2 k.k\varphi^\alpha) = 0.$$

Remarks.

1. The conservation law obtained for $\nu^2 k.k$ implies the conservation law for $\mathcal{E}$

$$\underline{\nabla}_\alpha(\varphi^\alpha \mathcal{E}) = 0.$$

This condition is in fact necessary for the coherence of the equations required from the background $\underline{g}_{\alpha\beta}$ and $\underline{\rho}_{\alpha\beta}$.

2. The equation $(4)_0$ is in fact deduced, as could have been foreseen from other considerations[8,9,12], from the equation

$$\varphi_0^2 R_{ij}^{(0)} - \varphi_0 \varphi_{(i} R_{j)0}^{(0)} + \varphi_i \varphi_j R_{00}^{(0)} = \varphi_0^2 \underline{\rho}_{ij} - \varphi_0 \varphi_{(i} \underline{\rho}_{j)0} + \varphi_i \varphi_j \underline{\rho}_{00}$$

Dominant Gravitational Wave, Propagation of the Fluid Wave.

The validity at order ω^{-1} of the fluid equations (5), (6) and (7) reads as a linear nonhomogeneous system in $(w'^\alpha, \psi', \zeta') = (\mathrm{w}'^I)$:

$$D_I^J \mathrm{w}'^I = B^J$$

with

$$B^\beta \equiv (\underline{e} + \underline{p})\{v^\alpha(\underline{\nabla}_\alpha \underline{u}^\beta + \gamma^\beta_{\alpha\lambda}\underline{u}^\lambda + \varphi_\alpha v'^\beta) + \underline{u}^\alpha(\underline{\nabla}_\alpha v^\beta + \gamma^\beta_{\alpha\lambda} v^\lambda + \gamma^{(1)\beta}_{\alpha\lambda}\underline{u}^\lambda)\} +$$
$$(\underline{f}'_p + 1)\pi(\underline{u}^\alpha \underline{\nabla}_\alpha \underline{u}^\beta + \underline{u}^\alpha \underline{u}^\lambda \gamma^\beta_{\alpha\lambda} + a v'^\beta) + (\underline{g}^{\alpha\beta} + \underline{u}^\alpha \underline{u}^\beta)\partial_\alpha \pi +$$
$$(\underline{u}^\alpha v^\beta + \underline{u}^\beta v^\alpha - h^{\alpha\beta})(\underline{\partial}_\alpha \underline{p} + \varphi_\alpha \pi'),$$

$$B^4 \equiv \underline{f}'_p\{\underline{u}^\alpha \underline{\partial}_\alpha \pi + v^\alpha(\underline{\partial}_\alpha \underline{p} + \varphi_\alpha \pi')\} + \underline{f}''_{pp}\pi \underline{u}^\alpha(\underline{\partial}_\alpha \underline{p} + \varphi_\alpha \pi') +$$
$$(\underline{f}'_p + 1)\pi(\underline{\nabla}_\alpha \underline{u}^\alpha + \gamma^\alpha_{\alpha\lambda}\underline{u}^\lambda + \varphi_\alpha v'^\alpha) + (\underline{e} + \underline{p})\underline{u}^\lambda \gamma^{(1)\alpha}_{\alpha\lambda} +$$
$$(\underline{e} + \underline{p})(\underline{\nabla}_\alpha v^\alpha + \gamma^\alpha_{\alpha\lambda} v^\lambda),$$

$$B^5 \equiv -w^\alpha \partial_\alpha \underline{S}.$$

Case (a), $\underline{f}'_p \neq 1$. Then $\mathrm{D} \neq 0$, w', π', ζ' are determined linear functions of the B^I, with coefficients independent of ξ. The expansion will have asymptotic validity if and only if these linear functions $\mathrm{F}(\mathrm{x},\xi)$ satisfy the asymptotic condition (zero mean in ξ): a sraightforward computation shows that this condition is satisfied.

Case (b), $\underline{f}'_p = 1$. The linear system in w', π', ζ' has a solution if and only if its non homogeneous term B^I is such that

$$B^I \ell_I = 0.$$

By replacing v^I by its value $d^I U + V^I$ we see that this equation is a propagation equation for U, analogous to the one found in[3] in the case of a fixed metric, but it is now non homogeneous: the gravitational wave generates a fluid wave. We find, by using the values of d, ℓ, and V^I together with the algebraic conditions, that this equation reads

$$\varphi^\alpha \underline{\partial}_\alpha U + \Phi U + \Psi U U' = \mathcal{P}$$

where Φ and Ψ are functions of x depending only on the background and φ, identical to the one found in [3] for a fixed metric, in particular

$$\Psi \equiv 0$$

if $\underline{f}''_{pp} = 0$.

The source term $\mathcal{P}$ is determined by the gravitational wave. It is a linear function of k, given by

$$\gamma^\lambda_{\alpha\beta} \underline{\nabla}_\lambda (\underline{u}^\alpha \underline{u}^\beta)$$

Remark.It was already known[3] that there is no distorsion of signals for fluid waves in relativistic "hard " fluids: there is no term in UU' (fluid waves in an incompressible fluid are exceptionnal in the sense of Lax - Boillat). We find moreover that the propagation equations are linear in k, and contain neither products of k or its derivatives with U or its derivatives, nor χ_{ij}. In these computations we use in particular the property, which holds if $\varphi_\alpha \varphi^\alpha = 0$ and $\varphi^i k_{ij} = 0$, that

$$\varphi_\beta \gamma^\beta_{\alpha\lambda} = 0, \quad \varphi_\beta V^\beta = 0$$

Conclusion and Open Problems.

We have shown that high frequency gravitational and fluid waves both of order one can be an asymptotic solution of order one of the Einstein equations, and of order two of the fluid equations if and only if the phase φ is isotropic. The gravitational wave is independent of the fluid wave , the fluid wave depends on the gravitational wave if

$$\underline{u}^i k_{ij} \not\equiv 0 \text{ for some j, or } \mathcal{P} \not\equiv 0$$

We have given the algebraic form of the wave and its propagation equations along the rays φ^α.

The interaction of gravitational and fluid waves of the same phase but other relative strengths can be studied by the same method. We will give the results elsewhere.

The case of gravitational and fluid waves of different phases is much more difficult, as it is well known for general quasi linear P.D.E. The interaction of high frequency waves of different phases can be tackled in a simple way only for two equations on two dimensional space times [10]. Perhaps the results obtained by Hunter, Majda and Rosales[11] for the interaction of waves with three different phases satisfying special relations could lead to interesting properties for the interaction of gravitational and fluid waves.

References

[1] P. Lax Duke Math. J. **24** p.627-646, 1957.
[2] L. Garding, T. Kotake, J. Leray Bull. Soc. Math. Fr. **94** p.25-48, 1966
[3] Y. Choquet-Bruhat J. Math. Pures et App. **48** p.117-158, 1969
[4] Y. Choquet-Bruhat in Relativity and Gravitation, Kuper and Peres ed. Gordon and Breach 1971
[5] M. Anile Relativistic fluids and magnetofluids, Cam. Univ. Press, 1989
[6] Y. Choquet-Bruhat and A. Taub J. of Gen. Rel. and Grav **8** p.561-571, 1977
[7] Y. Choquet-Bruhat and A. Greco J. Math. Phys. **24** p.377-380, 1983
[8] Y. Choquet-Bruhat and T. Ruggeri Comm. Maths. Phys. **89** p.269-275, 1983
[9] Y. Choquet-Bruhat and J.W. York C.R.Acad. Sci. Paris **321** p.1089-1095, 1995
[10] Y. Choquet-Bruhat C.R. Acad. Sci. Paris **268** p.1560-1563 1971
[11] J.K. Hunter, A. Majda and R. Rosales Studies in App. Math. **75** p.187-226, 1986
[12] A. Abrahams, A. Anderson, Y. Choquet-Bruhat and J.W. York, in preparation
[13] J. Ehlers, A.R. Prasanna and R. Breuer, Class. Quant. Grav. **4** p.253-264, 1987.

This page intentionally left blank

COURSE 6

POSITIVE-ENERGY THEOREMS

Yvonne CHOQUET-BRUHAT

Département de Mécanique
Faculté des Sciences, Université de Paris
Paris, France

B.S. DeWitt and R. Stora, eds.
Les Houches, Session XL, 1983
Relativité, groupes et topologie II / Relativity, groups and topology II
© *Elsevier Science Publishers B.V., 1984*



Contents

Introduction 742
1. Definitions 743
  1.1. Stress-energy $\hat{g}$-tensor 743
  1.2. Energy-momentum $\hat{g}$ vector 744
  1.3. Asymptotically Minkowskian manifolds 747
    1.3.1. ADM energy-momentum 751
  1.4. Hamiltonian formulation 753
2. Positivity of mass in a neighborhood of Minkowski space 758
  2.1. Functional spaces and isomorphism theorem 758
  2.2. Time-symmetric initial data sets as minima of energy 759
  2.3. Existence of maximal slices 760
  2.4. The mass functional and its derivatives 762
3. The proof of Schoen and Yau 766
  3.1. Second variation of the area 766
  3.2. Metrics Euclidean outside of a compact set 767
  3.3. Asymptotically Euclidean manifolds 768
    3.3.1. Extension of the Schoen–Yau proof to the case $\mathrm{tr}\, k \neq 0$ 771
4. Witten's proof 772
  4.1. Definition of the Witten two-form 773
  4.2. Proof of positivity 775
  4.3. Solution of the Witten equation 777
  4.4. Zero energy implies flatness of space-time 778
5. Recent developments 779
  5.1. Positive-mass theorem for black holes 779
  5.2. The supergravity origin of Witten's expression 781
  5.3. Asymptotically anti-De Sitter space-times 783
References 784

Introduction

It is well known that in Einstein theory there is no intrinsic definition of a local density of gravitational energy. Physically this is a consequence of the equivalence principle (the gravitational field appears zero to an observer in free fall), mathematically it results from the fact that a connection is not a tensor. However when a background space–time is given in addition to the one we are studying, it may be possible to define an energy for the latter with respect to the former. For instance, a local positive density of energy can be defined for gravitational waves propagating in a fixed background (high-frequency waves, shock waves), by an approximation method or, better, by a rigorous asymptotic expansion. It is also possible to define the global energy of a space–time $(V, {}^{(4)}g)$ with respect to another, given, one $(V, {}^{(4)}\hat{g})$, provided the latter admits a time-like Killing vector field, and the two metrics are asymptotically the same at space-like infinity. It turns out that the energy can be expressed as a flux integral at spatial infinity—in accordance with the fact that the energy of a gravitational field plays also the usual role of the charge in gauge theories. If one takes as background the vacuum of classical General Relativity, Minkowski space–time $(\mathbb{R}^4, \eta)$ the global energy of an asymptotically Minkowskian space–time corresponds to our idea of an isolated gravitating system whose metric at large spatial distances tends to the Minkowski metric. In this context the energy is often called the mass of the gravitating system, it coincides with m for the Schwarzschild solution. It has been conjectured that this mass (energy) is positive for all space–times satisfying reasonable assumptions, in particular having sources with non-negative local energy density. This conjecture had been proved to be correct in many special cases along the years (for references see for instance the review by Brill and Pong 1980). The full positivity theorem for asymptotically Minkowskian spaces has only been proved recently, by methods of geometrical (Schoen and Yau 1979) or super-symmetric (Witten 1981) origin. It is one more remarkable rigorous mathematical result which supports the physical relevance of classical General Relativity.

One may add that in all physical theories the energy and its positivity play a fundamental role, already at the classical level, not to speak of the

construction of Hilbert spaces relevant to quantization. Indeed, a solution of a classical conservative dynamical system is stable if its energy is the strict minimum of the energies of all possible solutions. The fact that Minkowski space–time is the only one with zero "mass" (i.e. gravitational energy) is an indication that this space–time is stable—in some sense whose precise mathematical definition remains to be given. Unfortunately the fact that the gravitational energy considered is a non-local functional of the fields makes it difficult to use for the proof of global existence or stability of solutions of Einstein equations.

The purpose of these lectures is to review the recent rigorous mathematical proofs of the positivity of energy of asymptotically Minkowskian space–times. In § 1 we recall definitions of global gravitational energy, from a covariant (following Abbott and Deser 1982, in accordance with the general construction of conserved quantities in non-linear theories with invariance groups as given by DeWitt 1984) or a Hamiltonian point of view (following Regge and Teitelboim 1974). In § 2 we give the analysis of the energy functional and the proof of its positivity in a neighborhood of flat space (as in Choquet-Bruhat and Marsden 1976). In § 3 we indicate the main ideas of the proof of Schoen and Yau (1979). In § 4 we give a complete and self contained exposition of Witten's (1981) proof. In the last section, § 5, we give a brief account of various extensions which have been obtained recently, and the classical supergravity interpretation of Witten's proof. We have not included in these lectures—because of lack of space and time—related topics, like the positivity of the Bondi mass or the problem of the semi-classical stability (cf. Witten 1982 and references therein).

1. Definitions

1.1. Stress-energy $\hat{g}$-tensor

Let $(M, {}^{(4)}\hat{g})$ be a given hyperbolic space–time satisfying the empty-space Einstein equations, with possibly a cosmological term,

$$\hat{S}^{\alpha\beta} \equiv \hat{R}^{\alpha\beta} - \tfrac{1}{2}\hat{g}^{\alpha\beta}\hat{R} + \Lambda\hat{g}^{\alpha\beta} = 0, \quad \text{i.e. } S({}^{(4)}\hat{g}) = 0. \tag{1.1}$$

The stress-energy gravitational $\hat{g}$-tensor t of another hyperbolic space–time $(M, {}^{(4)}g)$ is defined as the difference between the linearization of $S({}^{(4)}g)$ around ${}^{(4)}\hat{g}$, and $S({}^{(4)}g)$ itself, namely t is the tensor on M:

$$t({}^{(4)}g, {}^{(4)}\hat{g}) \equiv S'({}^{(4)}\hat{g}) \cdot ({}^{(4)}g - {}^{(4)}\hat{g}) - S({}^{(4)}g), \tag{1.2}$$

where $S'({}^{(4)}\hat{g})$ is the derivative at the point ${}^{(4)}\hat{g}$ of the mapping ${}^{(4)}g \to S({}^{(4)}g)$. If ${}^{(4)}g$ is a solution of Einstein equations with cosmological constant Λ and matter sources the symmetric two-tensor T, that is if

$$S({}^{(4)}g) = T, \quad \text{i.e. } S^{\alpha\beta} \equiv R^{\alpha\beta} - \tfrac{1}{2}g^{\alpha\beta}R + \Lambda g^{\alpha\beta} = T^{\alpha\beta},$$

then the identity (1.2) gives the equation

$$S'({}^{(4)}\hat{g})h = t + T, \qquad h = {}^{(4)}g - {}^{(4)}\hat{g}, \quad \text{i.e. } h_{\alpha\beta} = g_{\alpha\beta} - \hat{g}_{\alpha\beta}. \tag{1.3}$$

The Bianchi identities

$$\nabla \cdot S({}^{(4)}g) \equiv 0, \quad \text{i.e. } \nabla_\alpha S^{\alpha\beta} \equiv 0,$$

imply the linearized Bianchi identities (remember $S({}^{(4)}\hat{g}) = 0$)

$$\hat{\nabla} \cdot (S'({}^{(4)}\hat{g})h) \equiv 0, \tag{1.4}$$

therefore eq. (1.3) implies the conservation law

$$\hat{\nabla}_\alpha(t^{\alpha\beta} + T^{\alpha\beta}) = 0. \tag{1.5}$$

We have, in local coordinates (Lichnerowicz 1961),

$$\begin{aligned}\{S'({}^{(4)}\hat{g}) \cdot h\}^{\alpha\beta} = \tfrac{1}{2}\{ &- \hat{\nabla}_\lambda \hat{\nabla}^\lambda H^{\alpha\beta} + \hat{\nabla}^\alpha \hat{\nabla}_\lambda H^{\beta\lambda} + \hat{\nabla}^\beta \hat{\nabla}_\lambda H^{\alpha\lambda} - \hat{g}^{\alpha\beta}\hat{\nabla}_\lambda \hat{\nabla}_\mu H^{\lambda\mu} \\ &+ \hat{R}^\alpha{}_\mu H^{\mu\beta} + \hat{R}^\beta{}_\mu H^{\mu\alpha} - 2\hat{R}_\mu{}^\alpha{}_\nu{}^\beta H^{\mu\nu} + H^{\lambda\mu}\hat{R}_{\lambda\mu}\hat{g}^{\alpha\beta} \\ &- \hat{R}H^{\alpha\beta} + 2\Lambda H^{\alpha\beta} - \Lambda \,\hat{\mathrm{tr}}\, H\, \hat{g}^{\alpha\beta}\Big\}\end{aligned} \tag{1.6}$$

with $H^{\alpha\beta} = h^{\alpha\beta} - \frac{1}{2}\hat{g}^{\alpha\beta}\,\hat{\mathrm{tr}}\, h, \qquad \hat{\mathrm{tr}}\, h = \hat{g}^{\lambda\mu}h_{\lambda\mu}.$

1.2. Energy–momentum $\hat{g}$ vector

Given an arbitrary vector field ξ on M, it associates to the stress energy $\hat{g}$-tensor $t+T$ a one-form $p(\xi)$:

$$p_\alpha = (t_{\alpha\beta} + T_{\alpha\beta})\xi^\beta. \tag{1.7}$$

From the conservation law (1.5) one deduces a conservation law (in the ${}^{(4)}\hat{g}$ metric) for p if ξ is a Killing vector field of ${}^{(4)}\hat{g}$. Indeed if

$$\hat{\nabla}_\alpha \xi_\beta + \hat{\nabla}_\beta \xi_\alpha = 0,$$

then

$$\hat{\nabla}_\alpha p^\alpha = 0. \tag{1.8}$$

If ξ is time-like, p (index raised with ${}^{(4)}\hat{g}$) is called the (local) energy–

momentum ${}^{(4)}\hat{g}$ vector relative to ξ. From the conservation law (1.8) one deduces the usual relation between variation of (total) $\hat{g}$-energy and flux of $\hat{g}$-momentum,

$$\int_{\partial D} p^{\alpha} n_{\alpha} \, \mathrm{d}(\partial D) = 0, \tag{1.9}$$

where $\mathrm{d}(\partial D)$ is the volume element of the metric induced by ${}^{(4)}\hat{g}$ on the three-dimensional boundary ∂D of a four-dimensional domain D of M, and n_{α} is the unit normal to ∂D.

We suppose that M is a product $\Sigma \times \mathbb{R}$, with each $\Sigma_t = \Sigma \times \{t\}$ space-like, and the lines $\{x\} \times \mathbb{R}$, $x \in \Sigma$, time-like. The local coordinates $(x^{\alpha}) = (x^i, x^0)$ are adapted to the product structure. We choose $D = \Sigma \times (0, t)$. Then (1.9) reads [$\hat{\alpha} = (\hat{g}^{00})^{-1/2}$ is the lapse function of ${}^{(4)}\hat{g}$]:

$$\int_{\Sigma_t} \hat{\alpha} p^0 \, \mathrm{d}\Sigma_t - \int_{\Sigma_0} \hat{\alpha} p^0 \, \mathrm{d}\Sigma_0 = \int_0^t \int_{\partial \Sigma_\tau} \hat{\alpha} p^i v_i \, \mathrm{d}(\partial \Sigma_\tau) \mathrm{d}\tau. \tag{1.10}$$

The integral over $\partial \Sigma_t$ must be understood as the limit as n tends to infinity of integrals over ∂K_n, where the K_n are a sequence of increasing compact sets whose union covers Σ. We have denoted by v the unit normal to $\partial \Sigma$ in Σ, and by $\mathrm{d}(\partial \Sigma)$ the volume element of $\partial \Sigma$. From (1.10) we deduce the classical formula relating the derivative of the energy to the flux of momentum:

$$\frac{\mathrm{d}}{\mathrm{d}t} E = \int_{\partial \Sigma_t} \hat{\alpha} p^i v_i \, \mathrm{d}(\partial \Sigma_t), \tag{1.11}$$

where E, the ${}^{(4)}\hat{g}$-energy relative to ξ and the slice Σ_t, is

$$E = \int_{\Sigma_t} \hat{\alpha} p^0 \, \mathrm{d}\Sigma_t. \tag{1.12}$$

When Σ_t, ${}^{(4)}\hat{g}$, ξ, ${}^{(4)}g$ are such that the flux of momentum in (1.11) is zero, then the energy E is independent of t. This conservation of energy holds in particular if Σ_t, endowed with the metric induced by ${}^{(4)}\hat{g}$, is a complete Riemannian manifold, and if p has the appropriate fall-off at infinity.

More generally, for an arbitrary slice Σ of M, the ${}^{(4)}\hat{g}$ energy relative to ξ and the slice Σ is (when this integral is defined)

$$E = \int_{\Sigma} p^{\alpha} n_{\alpha} \mathrm{d}\Sigma, \tag{1.12a}$$

which we can also write, introducing the one-form p and its adjoint $*p$ (indices lowered with ${}^{(4)}\hat{g}$, notations of Choquet-Bruhat et al. 1982),

$$p = p_{\alpha}\mathrm{d}x^{\alpha}, \quad *p = \frac{1}{3!}\hat{\tau}_{\alpha\lambda\mu\nu}p^{\alpha}\,\mathrm{d}x^{\lambda}\wedge\mathrm{d}x^{\mu}\wedge\mathrm{d}x^{\nu}, \quad E = \int_{\Sigma} *p. \tag{1.13}$$

Equation 1.8 expresses that

$$\delta p = 0,$$

thus there exists, at least locally, a two-form such that

$$p = \delta q, \qquad \text{equivalently} \quad *p = \mathrm{d}(*q).$$

In fact we shall prove

Lemma. *There exists (globally) on M a two-form q, called a superpotential, such that $p = \delta q$, and thus*

$$E = \int_{\partial\Sigma} *q. \tag{1.14}$$

Proof. We construct explicitly an antisymmetric two-tensor $q^{\mu\alpha}$ such that

$$p^{\mu} = \hat{\nabla}_{\alpha} q^{\mu\alpha}. \tag{1.15}$$

Indeed, we have, by definition [cf. eq. (1.3)]:

$$p^{\mu} = (t^{\mu\nu} + T^{\mu\nu})\xi_{\nu} = \{S'({}^{(4)}\hat{g})h\}^{\mu\nu}\xi_{\nu}. \tag{1.16}$$

We deduce from eq. (1.6) (cf. Abbott and Deser 1982),

$$\{S'({}^{(4)}\hat{g})h\}^{\mu\nu} \equiv \hat{\nabla}_{\alpha}\hat{\nabla}_{\beta}K^{\mu\alpha\nu\beta} + X^{\mu\nu}, \tag{1.17}$$

where $K^{\mu\alpha\nu\beta}$ is the four-tensor with the symmetries of the Riemann tensor:

$$K^{\mu\alpha\nu\beta} = \tfrac{1}{2}(\hat{g}^{\mu\beta}H^{\nu\alpha} + \hat{g}^{\nu\alpha}H^{\mu\beta} - \hat{g}^{\mu\nu}H^{\alpha\beta} - \hat{g}^{\alpha\beta}H^{\mu\nu}) \tag{1.18}$$

and

$$X^{\mu\nu} = \tfrac{1}{2}(\hat{R}^{\mu}{}_{\alpha\beta}{}^{\nu}H^{\alpha\beta} - \Lambda H^{\mu\nu}). \tag{1.19}$$

Using the equation $\hat{R}_{\alpha\beta} = \Lambda \hat{g}_{\alpha\beta}$ satisfied by the background and the definition (1.18) of $K^{\alpha\beta\lambda\mu}$ one shows that

$$X^{\mu\nu} = \tfrac{1}{2}\hat{R}^{\nu}{}_{\lambda\alpha\beta}K^{\mu\lambda\alpha\beta}.$$

From eqs. (1.16) and (1.17) we deduce

$$p^{\mu} = \hat{\nabla}_{\alpha}[(\hat{\nabla}_{\beta}K^{\mu\alpha\nu\beta})\xi_{\nu} - K^{\mu\beta\nu\alpha}\hat{\nabla}_{\beta}\xi_{\nu}] + K^{\mu\alpha\nu\beta}\hat{\nabla}_{\beta}\hat{\nabla}_{\alpha}\xi_{\nu} + X^{\mu\nu}\xi_{\nu}. \tag{1.20}$$

Since ξ^{ν} is a Killing vector it satisfies the identity

$$\hat{\nabla}_{\alpha}\xi_{\beta} + \hat{\nabla}_{\beta}\xi_{\alpha} = 0.$$

From this we deduce, using the Ricci identity and the algebraic Bianchi identity, that

$$\hat{\nabla}_{\beta}\hat{\nabla}_{\alpha}\xi_{\nu} = \tfrac{1}{2}\hat{R}_{\nu\alpha,\beta}{}^{\lambda}\xi_{\lambda};$$

thus we find, as announced,

$$p^{\mu} = \hat{\nabla}_{\alpha}q^{\mu\alpha},$$

with $q^{\mu\alpha}$ the antisymmetric two-tensor

$$q^{\mu\alpha} = (\hat{\nabla}_{\beta}K^{\mu\alpha\nu\beta})\xi_{\nu} - K^{\mu\beta\nu\alpha}\hat{\nabla}_{\beta}\xi_{\nu}. \tag{1.21}$$

The considered energy E is therefore given by an integral on the boundary of Σ:

$$E = \int_{\partial\Sigma} {}^{*}q = \int_{\partial\Sigma} \tfrac{1}{4}\hat{\tau}_{\lambda\beta\mu\alpha}q^{\mu\alpha}\,\mathrm{d}x^{\lambda} \wedge \mathrm{d}x^{\beta}. \tag{1.22}$$

In coordinates where Σ is $x^0 = \text{constant}$:

$$E = \int_{\partial\Sigma} q^{0i}\mathrm{d}S_i, \quad \text{with} \quad \mathrm{d}S_i = \tfrac{1}{2}(\det{}^{(4)}\hat{g})^{1/2}\varepsilon_{ijk}\mathrm{d}x^{j} \wedge \mathrm{d}x^{k}. \tag{1.23}$$

1.3. Asymptotically Minkowskian manifolds

Let us consider a gravitating isolated system, for instance the solar system. A natural space–time model for it is a manifold $M = \mathbb{R}^3 \times \mathbb{R}$ with a hyperbolic metric ${}^{(4)}g$ which outside of a subset $K \times \mathbb{R}$ of M (with K some large compact set) will look like the Minkowskian metric η, in rectangular coordinates if the coordinates in the asymptotic region $(\mathbb{R}^3 \setminus K) \times \mathbb{R}$ are "inertial"; a notion which can be physically defined by

the observation of distant galaxies. Thus there will be a possibility to define asymptotically Minkowskian metrics ${}^{(4)}g$ on M as those whose difference with η tends to zero, in a sense to explain in $(\mathbb{R}^3 \setminus K_n) \times \mathbb{R}$, when n tends to infinity and K_n is a sequence of increasing compact sets whose union covers $\mathbb{R}^3$. To allow for a more general definition of asymptotically Minkowskian space-times, with $M = \Sigma \times \mathbb{R}$, and Σ not necessarily diffeomorphic to $\mathbb{R}^3$, we introduce first the notion of a space-time "Minkowskian at infinity" as follows.

Definition 1. A manifold $M = \Sigma \times \mathbb{R}$ endowed with a hyperbolic metric $\hat{\eta}$ with $\Sigma \times \{t\}$ space-like, $\{x\} \times \mathbb{R}$ time-like is said to be Minkowskian at infinity if

(1) There is a compact set $K \subset \Sigma$ such that $\Sigma \setminus K$ is the disjoint union of a finite number of open sets Ω_A, $A = 1, \ldots, N$. Each Ω_A is diffeomorphic to the complement of a closed ball $\bar{B}_A$ of $\mathbb{R}^3$, in a diffeomorphism ϕ_A.

(2) In each $\Omega_A \times \mathbb{R}$ the metric $\hat{\eta}$ is flat; it is the Minkowski metric, $\mathrm{diag}(+1, +1, +1, -1)$ in the coordinates defined by ϕ_A.

The different subsets Ω_A are sometimes called the "ends" of Σ.

We can now define asymptotically Minkowskian space-times as follows:

Definition 2. A space-time $(M = \Sigma \times \mathbb{R}, {}^{(4)}g)$ is said to be asymptotically Minkowskian if

(1) Σ satisfies condition 1 of definition 1;

(2) In each Ω_A the difference ${}^{(4)}g - \hat{\eta}$ "tends to zero at infinity", in a sense to explain, for each $t \in \mathbb{R}$.

The two main definitions which have been used for the fall-off at infinity of gravitational fields have been first through the use of continuous functions and, more recently, by introducing weighted Sobolev spaces: they both capture, roughly, the same fall-off features, but not exactly. One or the other is more adapted to the mathematical treatment at hand. We recall that a space-time tensor, for instance ${}^{(4)}g - \hat{\eta}$, on $M = \Sigma \times \mathbb{R}$ admits a canonical decomposition into tensors on Σ (vectors, scalar) depending on $t \in \mathbb{R}$.

The norm $|f(x)|$ of a tensor f at a point $x \in \Sigma$ and its derivative ∂f are defined with the properly Riemannian metric $\hat{e}$ induced by $\hat{\eta}$ on Σ (in Ω_A, $\hat{e}$ is the Euclidean metric, $\sum (dx^i)^2$ in ϕ_A coordinates). We introduce a function σ on Σ, positive and smooth, which characterizes the distance to a fixed, arbitrary point. Namely we impose to σ that,

for $x \in \Omega_A$, in ϕ_A coordinates

$$\sigma(x) = (1+r^2)^{1/2}.$$

We define the following functional spaces of tensor fields (of some given type) on Σ.

C^s_δ: tensor fields on Σ which are s times continuously differentiable and such that

$$\|f\|_{C^s_\delta} \equiv \sum_{k=0}^{s} \operatorname{Sup}_{x \in \Sigma} \sigma(x)^{\delta+k} |\partial^k f(x)| < +\infty ;$$

H^s_δ: tensor fields on Σ whose generalized (in the sense of distribution theory) derivations of order $\leqslant s$ are functions such that $\sigma^{\delta+k}\partial^k f$ is square integrable on Σ (volume element of $\hat{e}$). H^s_δ is a Hilbert space for the norm

$$\|f\|_{H^s_\delta} = \left\{ \int_\Sigma \sum_{k=0}^{s} \sigma^{2\delta+2k} \partial^k f|^2 \, \mathrm{d}\Sigma \right\}^{1/2}.$$

Remark. Σ is a complete Riemannian manifold Euclidean at infinity, therefore the space $\mathscr{D}$ (i.e. C^∞ with compact support) is dense in H^s_δ for every $s \in \mathbb{N}$, $\delta \in \mathbb{R}$.

The property $f \in C^s_\delta$ is equivalent to the classical notation

$$f = O(r^{-\delta}), \qquad \partial f = O(r^{-\delta-1}), \qquad \partial^s f = O(r^{-\delta-s}).$$

The property $f \in H^s_\delta$ does not have so obvious a meaning, but the following inclusion theorem sheds some light on it.

Theorem (Choquet-Bruhat and Christodoulou 1981). *On an n-dimensional manifold Euclidean at infinity the following inclusions hold*:

$$C^s_{\delta_1} \subset H^s_\delta \subset C^{s'}_{\delta'}, \quad \text{if } \delta' < \delta + n/2 < \delta_1, \quad s > s' + n/2.$$

(The first inclusion is easy to check; the second is more subtle.) For the case $n = 3$ the property $f \in H^s_\delta$ will insure that f tends to zero in a classical sense as r tends to infinity if $\delta + \frac{3}{2} > 0$, i.e. $\delta > -\frac{3}{2}$. If we also want that the classical fall-off property $f = O(r^{-1})$ insures that $f \in H^s_\delta$, then we must take $\delta < -\frac{1}{2}$. A possible choice in order to have $-\frac{3}{2} < \delta < -\frac{1}{2}$ is $\delta = -1$. The space H^1_{-1} has the interesting property that ∂f is square integrable. The quantiy

$$\int_\Sigma |\partial f|^2 \mathrm{d}\Sigma$$

is the (mathematical—and sometimes physical) energy of the field f.

We also have the obvious, but interesting property that if $f \in H^0_\delta$ and $h \in H^0_{\delta+2}$, and $\delta = -1$, fh is integrable.

Another useful theorem is

Theorem. *The continuous multiplication properties hold by product of tensor fields* $(f, g) \to f \otimes g$ *on an n-dimensional manifold* Σ *Euclidean at infinity*:

(1) $C^s_{\delta_1} \times C^s_{\delta_2} \to C^s_{\delta_1+\delta_2} \quad \forall s \in \mathbb{N}, \quad \delta_1, \delta_2 \in \mathbb{R}$;

(2) $H^{s_1}_{\delta_1} \times H^{s_2}_{\delta_2} \to H^s_\delta$ if $s \leqq s_1, s_2,\ s < s_1 + s_2 - n/2,\ \delta < \delta_1 + \delta_2 + n/2$.

H^s_δ *is a Banach algebra if* $s > n/2$, $\delta > -n/2$.

The proof of (1) is trivial. For the proof of (2) see Choquet-Bruhat and Christodoulou (1981).

The classical *definition of an asymptotically Euclidean initial data set* (g, k) on Σ is

$$g - \hat{e} \in C^s_1, \qquad k \in C^{s-1}_2;$$

i.e.,

$$g_{ij} - \delta_{ij} = O(r^{-1}), \text{---}, \partial^s g_{ij} = O(r^{-1-s}),$$
$$k = O(r^{-2}), \text{---}, \partial^{s-1} k = O(r^{-1-s}).$$

An alternative definition, more convenient in functional analysis, is

$$g - \hat{e} \in H^s_\delta, \qquad k \in H^{s-1}_{\delta+1}$$

for some δ with $-\frac{3}{2} < \delta < -\frac{1}{2}, \qquad s \geqslant 2$.

Such an initial data set is said to be H^s_δ asymptotically Euclidean.

We shall give in the same spirit the *definition of an asymptotically Minkowskian metric* ${}^{(4)}g$ on $\Sigma \times I$, I closed interval of $\mathbb{R}$. As it is well known the metric ${}^{(4)}g$ is determined by the metric g it induces on Σ, the projection of the vector tangent to the time-like line $\{x\} \times \mathbb{R}$ on Σ (shift vector β) and the normal to Σ (lapse function α). In coordinates:

$$g_{\alpha\beta} dx^\alpha du^\beta \equiv g_{ij} dx^i dx^j + 2\beta_i dx^i dt - (\alpha^2 - \beta^i \beta_i) dt^2$$

(note also $\alpha^2 = (-g^{00})^{-1}$).

We denote by

$$E^s_\delta = \bigcap_{0 \leqslant k \leqslant s} C^k(I, H^{s-k}_{\delta+k})$$

a (Banach) space of tensor fields f_t on Σ depending on $t \in I$; such that $\partial^k f_t/\partial t^k \in H^{s-k}_{\delta+k}$ for each t, and the mapping $I \to H_{s-k}$ by $t \to \partial^k f/\partial t^k$ is continuous and bounded.

The norm in E^s_δ is

$$\operatorname*{Sup}_{t\in I} \sum_{k=0}^{s} \|\partial^k f/\partial t^k\|_{H^s_\delta} = \|f\|_{E^s_\delta}.$$

Definition 3. A metric ${}^{(4)}g$ on $\Sigma \times I$ is said to be E^s_δ asymptotically Minkowskian if

$$g - \hat{e} \in E^s_\delta, \qquad \alpha - 1 \in E^s_\delta, \qquad \beta \in E^s_\delta.$$

We say simply that ${}^{(4)}g$ is H^s_δ-asymptotically Minkowskian on Σ_t if for $k \leqslant s$, $\partial^k(g-\hat{e})/\partial t^k$, $\partial^k(\alpha-1)/\partial t^k$ and $\partial^k\beta/\partial t^k$ are in H^s_δ at time t. (For the preservation of these hypotheses with time, cf. Christodoulou and O'Murchadha 1981.)

1.3.1. ADM energy-momentum

By eq. (1.12) the $\hat{\eta}$-energy relative to a time-like Killing vector ξ of $\hat{\eta}$, generator of translations in Ω_A, of an asymptotically Minkowskian manifold $(M, {}^{(4)}g)$ is well defined at time t if on Σ_t the metric ${}^{(4)}g$ is H^2_{-1} asymptotically Minkowskian, and the sources $T_{\alpha\beta}$ are integrable. To check this property one must go back to the definition (1.7) of p_α and use the fact that $t_{\alpha\beta}$ is an analytic function in the $g_{\lambda\mu}$, and a homogeneous quadratic form in $\partial_\nu g_{\lambda\mu}$ [the integrability of the three-form *p is no more visible when one uses (1.20), if one does not remember that g must satisfy the Einstein equation (1.16)].

We give below the standard formulae for the ADM energy and momentum—keeping in mind that they are meaningful only for a solution of Einstein equations with integrable sources.

We denote ${}^{(4)}g - \hat{\eta} = h$, i.e. in Ω_A

$$g_{\alpha\beta} - \eta_{\alpha\beta} = h_{\alpha\beta}.$$

We take coordinates (x^i, x^0) adapted to the product $\Sigma \times \mathbb{R}$, and $\hat{\eta}$-rectangular in $\Omega_A \times \mathbb{R}$ (then $\hat{\eta}_{\alpha\beta} = \eta_{\alpha\beta} = \operatorname{diag}(+1, +1, +1, -1)$; the formula (1.22) reads

$$P(\xi) = \sum_{A=1}^{N} \int_{S_A} q^{0k} \mathrm{d}S_k, \qquad \mathrm{d}S_k = \tfrac{1}{2}\varepsilon_{ijk}\mathrm{d}x^i \wedge \mathrm{d}x^j,$$

where the integral over S_A is the limit of the integral over a two-sphere in Ω_A whose radius tends to infinity. If we take coordinates $\hat{\eta}$-rectangular in the products $\Omega_A \times \mathbb{R}$, and denote by $U_{\lambda,A}$ the (constant) components of ξ in these coordinates, we obtain

$$P(\xi) \equiv \sum_{A=1}^{N} P_A^{\lambda} U_{\lambda,A} = \sum_{A=1}^{N} U_{\lambda,A} \int_{S_A} (-\eta^{\sigma k}\partial_\sigma H^{\lambda 0} + \eta^{\lambda k}\partial_\sigma H^{\sigma 0} + \eta^{\sigma 0}\partial_\sigma H^{\lambda k} - \eta^{\lambda 0}\partial_\sigma H^{\sigma k})\mathrm{d}S_k,$$

which gives, by a straightforward calculation, the standard ADM expressions (in rectangular coordinates in Ω_A)

$$P_A^0 = \int_{S_A} \partial_k H^{lk}\mathrm{d}S_k = \int_{S_A} (\partial_l g_{lk} - \partial_k g_{ll})\mathrm{d}S_j \tag{1.24}$$

and

$$P_A^l = \int_{S_A} (-\partial_k H^{l0} + \delta_k^l \partial_i H^{i0} - \partial_0 H^{lk})\mathrm{d}S_k,$$

which can also be written, by using the extrinsic curvature k of Σ in $(M, {}^{(4)}g)$, i.e. $k_{ij} = -\alpha\Gamma_{ij}^0$, and the fact that $\int_{\partial\Sigma}\partial_k f \mathrm{d}S_k = 0$,

$$P_A^l = -2\int_{S_A} (\delta^{lj}\mathrm{tr}\, k - k^{lj})\mathrm{d}S_j. \tag{1.25}$$

Remark 1. If Σ is not diffeomorphic to $\mathbb{R}^3$ (for instance one end and K not diffeomorphic to a closed ball of $\mathbb{R}^3$) it may not be possible to endow it with a smooth metric $\hat{e}$ such that $\hat{e} = e$ in the ends Ω_A (Euclidean metric in Ω_A coordinates) and the Killing field U (which generates translations in $\Omega_A \times \mathbb{R}$) extends to a Killing field in $\Sigma \times \mathbb{R}$. It is not clear then if the formula (1.22) can be interpreted as giving really a total energy-momentum on Σ of the asymptotically Minkowskian space-time $(M, {}^{(4)}g)$.

Remark 2. Suppose that $K = \Sigma \setminus \Omega$ (with $\Omega = \cup\, \Omega_A$, Ω_A diffeomorphic to $\mathbb{R}^3 \setminus \bar{B}_A$) is not compact, but we can complete it by a boundary ∂K such that $K \cup \partial K$ is compact. The ADM energy-momentum must then be

completed by

$$\int_{\partial K} q^{0k} \mathrm{d}S_k.$$

A case of particular interest is when Σ contains a black hole with apparent horizon ∂K (cf. § 5).

Example. The differentiable manifold $\mathbb{R}^3 \times \mathbb{R}$ endowed with a smooth metric ${}^{(4)}g$ defined by

$$\mathrm{d}s^2 = \left(1 + \frac{m}{2r}\right)^2 \sum_{i=1}^{3} (\mathrm{d}x^i)^2 - \left(\frac{1 - m/2r}{1 + m/2r}\right)^2 \mathrm{d}t^2 \quad \text{for } r > R > 2m$$

(Schwarzschild metric in isotropic coordinates) and a regular hyperbolic metric for $r \leqslant R$. This space-time is asymptotically Minkowskian ($\hat{\eta} = \eta = \mathrm{diag}(+1, +1, +1, -1)$ in the coordinates x^α). Its ADM energy-momentum is

$$P^0 = 16\pi m, \qquad P^l = 0.$$

It is a consequence of the positive-energy theorem which we shall prove later that $m \geqslant 0$ for realistic sources: it means that the Schwarzschild solution with $m < 0$ cannot be extended smoothly for $r \leqslant R$ to a solution of Einstein's equations with such sources.

The positive-energy theorem asserts that the ADM energy-momentum vector is time-like and future directed, zero only for Minkowski space-time itself. One defines the ADM mass by

$$16\pi m = |P^\lambda P_\lambda|^{1/2}.$$

1.4. Hamiltonian formulation

It is well known that Einstein equations in empty space can be formulated as a Hamiltonian system with constraints. Indeed, in the 3 + 1 formulation, for a space-time $(\Sigma \times \mathbb{R}, {}^{(4)}g)$, the fundamental unknown are, on each Σ_t, the metric g induced by ${}^{(4)}g$ and the extrinsic curvature k. The lapse α and shift β which together with g determine ${}^{(4)}g$ by

$$g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta = g_{ij}\mathrm{d}x^i \mathrm{d}x^j + 2\beta_i \mathrm{d}x^i \mathrm{d}t - (\alpha^2 - \beta^i\beta_i)\mathrm{d}t^2 \tag{1.26}$$

are gauge variables which can be adjusted by diffeomorphisms. The Einstein equations split into

(1) the constraints

$$\mathscr{H} \equiv -R + k \cdot k - (\operatorname{tr} k)^2 = 0, \tag{1.27a}$$

$$\mathscr{J} \equiv -2D \cdot (g \operatorname{tr} k - k) = 0 \quad [\text{i.e., } D_j(k_i{}^j - \delta_i{}^j k_l^l) = 0] \tag{1.27b}$$

(D covariant derivative in g);

(2) the evolution equations

$$G^{ij}({}^{(4)}g) = 0. \tag{1.28}$$

The data on Σ_0 of g and k satisfying the constraints determine the space–time (in its domain of global hyperbolicity with Σ_0 a Cauchy surface) up to isometries (for a new proof of this fact, which does not use harmonic coordinates, cf. Choquet-Bruhat and Ruggeri 1983). Moreover, the evolution equations (1.28) can be written as a Hamiltonian system for the conjugate variables g and $P = g \operatorname{tr} k - k$:

$$P^{ij} = g^{ij} \operatorname{tr} k - k^{ij},$$

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = \frac{\delta H}{\delta P^{ij}}, \qquad \frac{\mathrm{d}P^{ij}}{\mathrm{d}t} = -\frac{\delta H}{\delta g_{ij}}, \tag{1.29}$$

where the right-hand side denotes the L^2 functional partial derivative of the Hamiltonian H, a function on the phase space $(g, P) \mapsto H(g, P)$. Such derivatives are defined by the fact that for each allowed variation of g and P one has, at first order,

$$\delta H = \left(\frac{\delta H}{\delta g_{ij}}, \delta g_{ij}\right) + \left(\frac{\delta H}{\delta P^{ij}}, \delta P^{ij}\right), \tag{1.30}$$

where the (,) are L^2 scalar products,

$$(f, h) = \int_\Sigma f \cdot h \, \mathrm{d}\Sigma.$$

It has been remarked by DeWitt (1967) and exploited by Regge and Teitelboim (1974) that if one takes for H the ADM Hamiltonian

$$H_{\mathrm{ADM}} = \int_\Sigma (\alpha \mathscr{H} + \beta \cdot \mathscr{J}) \mathrm{d}\Sigma, \tag{1.31}$$

eq. (1.29) will not coincide with the actual Einstein equation (1.28), for a non-compact Σ, due to the appearance in δH of integrals on $\partial\Sigma$ which do not vanish for general variations δg, δk compatible with the asymptotic behavior: such will be the case if we consider asymptotically Euclidean

data g, k, on $(\Sigma, \hat{e})$ Euclidean at infinity, characterized by $g-\hat{e}\in H^2_\delta$, $k\in H^1_{\delta+1}$, with some δ such that $-\frac{3}{2}<\delta<-\frac{1}{2}$, or by $g-\hat{e}\in C^2_1$, $k\in C^1_2$.

Indeed, it can be shown (see references in Choquet-Bruhat and York 1980 or Fisher and Marsden 1979) that eq. (1.28), together with the classical expression for k,

$$\frac{\partial g_{ij}}{\partial t} = -2\alpha k_{ij}+D_i\beta_j+D_j\beta_i, \tag{1.32}$$

and the definition

$$P^{ij} = g^{ij}\operatorname{tr} k - k^{ij}, \tag{1.33}$$

can be written in a compact form

$$\begin{pmatrix} \mathrm{d}g/\mathrm{d}t \\ \mathrm{d}P/\mathrm{d}t \end{pmatrix} = J\circ(D\phi)^*(\alpha,\beta), \tag{1.34}$$

where J is the symplectic matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{1.35}$$

and $(D\phi)^*$ the L^2-formal adjoint of the linear differential operator which is the derivative at (g,P) of the mapping

$$\begin{pmatrix} g \\ P \end{pmatrix} \mapsto \phi(g;P) \equiv \begin{pmatrix} \mathscr{H}(g,P) \\ \mathscr{J}(g,P) \end{pmatrix}. \tag{1.36}$$

$D\phi$, at a point g, P, is a differential operator which maps a pair of symmetric two-tensor fields δg, δP on Σ into a scalar and a vector field on Σ. The L^2-formal adjoint is defined by

$$\int_\Sigma u\cdot(D\phi)v\,\mathrm{d}\Sigma \equiv \int_\Sigma v\cdot(D\phi)^*u\,\mathrm{d}\Sigma, \tag{1.37}$$

where u is a couple of a scalar and a vector and v a pair of symmetric two-tensors, C^2 with compact support on Σ.

On the other hand, we deduce from eq. (1.31) that, since

$$H_{\mathrm{ADM}} = \int_\Sigma \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\mathscr{H},\mathscr{J})\mathrm{d}\Sigma, \tag{1.38}$$

the functional derivative of H_{ADM} is

$$DH_{\mathrm{ADM}}\begin{pmatrix} \delta g \\ \delta P \end{pmatrix} = \int_\Sigma \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\cdot(D\phi)(\delta g,\delta P)\mathrm{d}\Sigma, \tag{1.39}$$

and we do not have

$$DH_{\mathrm{ADM}}\begin{pmatrix}\delta g\\ \delta P\end{pmatrix} = \int_{\Sigma}\begin{pmatrix}\delta g\\ \delta P\end{pmatrix}\cdot (D\phi)^{*}(\alpha,\beta)\mathrm{d}\Sigma \tag{1.40}$$

[which will be equivalent to the fact that eq. (1.34), the calculated expression for eq. (1.28), is of the Hamiltonian form (1.29)], because δg, δP are not necessarily with compact support. The expressions (1.39) and (1.40) are equal modulo an integral on the boundary $\partial\Sigma$. For an asymptotically Minkowskian space–time this integral reduces to (we suppress the summation on the several possible ends to simplify the writing)

$$-\int_{\partial\Sigma}(D_j(\delta g^{ij}) - D^i(\mathrm{tr}\,\delta g))\mathrm{d}S_l$$

which we recognize as the variation of

$$-\tilde{H} = \int_{\partial\Sigma}(g^{ij}\tilde{\Gamma}^l_{ij} - g^{jl}\tilde{\Gamma}^i_{ji})\mathrm{d}S_l,$$

where $\tilde{\Gamma}^l_{ij}$ denotes the difference of the connection of g and $\hat{e}$. Therefore if we replace the ADM Hamiltonian by the Regge–Teitelboim Hamiltonian,

$$H_{\mathrm{RT}} = H_{\mathrm{ADM}} + \tilde{H},$$

we obtain for (1.28) the correct Hamiltonian equations (1.29). In rectangular coordinates $\tilde{H}$ reads

$$\tilde{H} = \int_{\partial\Sigma}(\partial_i g_{il} - \partial_l g_{ii})\mathrm{d}S_l.$$

We verify that for a solution of the constraints the numerical value of the true Hamiltonian, generator of time translations in the sense (1.29), of an asymptotically Minkowskian space–time is indeed the *energy* we have calculated in the previous paragraph. The Hamilton equations (1.29) together with the constraints (1.27) show that it is conserved in time.

The definition of the *linear momentum* of an asymptotically Minkowskian space–time, and its conservation by time evolution, comes from the classical Noether theorem. Consider a one-parameter group of diffeomorphisms of Σ generated by a vector field a which reduces in Ω to a constant vector (i.e. $\partial a = 0$). The corresponding conserved quantity is

(L is the lie derivative, $\mathscr{L}$ the Lagrangian)

$$P(a) = -\left(\frac{\delta \mathscr{L}}{\partial \dot{g}_{ij}}, L_a g_{ij}\right) = -\int_\Sigma P^{ij}(\nabla_i a_j + \nabla_j a_i)\mathrm{d}\Sigma$$

$$= -2\int_{\partial\Sigma} P^{ij} a_j \,\mathrm{d}S_i + 2\int_\Sigma a_j \nabla_i P^{ij}\mathrm{d}\Sigma.$$

The second integral vanishes by the constraint (1.27b) and we are left with

$$P(a) = P^i a_i,$$

where

$$P^i = -2\int_{\partial\Sigma} P^{ij}\mathrm{d}S_j$$

is the linear momentum we have computed before.

Remark. P_k is the generator of space translations in the sense that, if Σ is diffeomorphic to $\mathbb{R}^3$ and $\partial a = 0$ on $\mathbb{R}^3$ then

$$P(a) = \int_\Sigma P^{ij} a^k \partial_k g_{ij}\,\mathrm{d}\Sigma;$$

thus

$$P_k = -\int_\Sigma P^{ij}\partial_k g_{ij}\,\mathrm{d}\Sigma$$

and then

$$\delta P_k = -\int_\Sigma (\delta P^{ij}\partial_k g_{ij} + P^{ij}\partial_k(\delta g_{ij}))\mathrm{d}\Sigma.$$

Thus, for any asymptotically Minkowskian space–time:

$$\delta P_k = -\int_\Sigma (\delta P^{ij}\partial_k g_{ij} - \delta g_{ij}\partial_k P^{ij})\mathrm{d}\Sigma,$$

that is,

$$-\partial_k g_{ij} = \frac{\delta P_k}{\delta P_{ij}}, \qquad -\partial_k P^{ij} = -\frac{\delta P_k}{\delta g_{ij}}.$$

2. Positivity of mass in a neighborhood of Minkowski space

The method used by Choquet-Bruhat and Marsden (1976) to prove the positivity of the mass of space–times which are, in a functional sense, near to Minkowski space–time, originates in ideas and partial results of Brill and Deser (1968). These authors determined for asymptotically flat spaces the first variation of the mass functional and found that it has only one critical point, namely at flat space. They then showed that the second variation at that point was positive, when restricted to physically meaningful variations of the metric. Our argument to obtain a rigorous and non-infinitesimal proof consisted essentially of two parts: first reduce the problem to a non-degenerate one by constructing a slice for the action of the group of isometries of the initial manifold; second, make the appropriate estimates to prove that the second derivative of the mass functional is indeed positive on this slice, at least in a finite neighborhood of its critical point.

2.1. Functional spaces and isomorphism theorem

The functional spaces we used were the $W^p_{s,\delta}$ weighted Sobolev-spaces of Nirenberg and Walker, studied by M. Cantor who has proved, among other things, that the Laplacian Δ_g of an asymptotically Euclidean metric on $\mathbb{R}^3$ is an isomorphism (bijective, continuous as well as its inverse, and of course linear mapping) $W^p_{s,\delta} \to W^p_{s-2,\delta+2}$, if $\delta = 0$, $p > 3$. In fact this property is true, more generally, if $p > 1$, and δ is appropriately chosen. It is interesting to choose $p = 2$, since the spaces $W^2_{s,\delta}$ denoted H^s_δ (cf. § 1) are Hilbert spaces. The theorem also holds for a general Σ, not necessarily $\mathbb{R}^3$. It says precisely*:

Theorem. *The operator $\Delta_g + f$ on functions over Σ, where Δ_g is the Laplace operator of an asymptotically Euclidean metric g on the n-dimensional*

* For a general study of elliptic systems on asymptotically Euclidean manifolds, cf. Choquet-Bruhat and Christodoulou (1981); for previous results using Holder spaces with weight, cf. Chaljub-Simon and Choquet-Bruhat (1979).

($n > 1$) manifold (Σ, e) Euclidean at infinity, is an isomorphism

$$H_{s,\delta} \to H_{s-2,\delta+2}, \quad s > n/2, \quad -n/2 < \delta < -2+n/2$$

if $f \leqslant 0$ and if

$$g-e \in H_{\sigma,\rho}, \quad \sigma > 1+n/2, \quad \rho > -n/2,$$

$$f \in H_{s_0,\delta_0}, \quad s_0 > 1+n/2, \quad \delta_0 > 2-n/2,$$

with $\inf(s_0, \sigma) \geqslant s-2$.

2.2. Time-symmetric initial data sets as minima of energy

The Hamiltonian constraint reads, in the presence of sources:

$$R - k \cdot k + (\operatorname{tr} k)^2 = 2\mu, \quad \mu = T^{00}. \tag{2.1}$$

If the sources satisfy the weak energy condition (cf. Hawking and Ellis 1973) we have $\mu \geqslant 0$. Thus eq. (2.1) will give

$$R \geqslant 0 \text{ on } \Sigma \quad \text{if } \operatorname{tr} k = 0 \text{ on } \Sigma. \tag{2.2}$$

The inequality (2.2), together with the previous theorem, will give a very simple proof of the following (we return to the physical case $n = 3$).

Theorem (O'Murchadha and York 1974). *For each initial data set (g, k) on (Σ, e), with $\operatorname{tr} k = 0$, for an Einstein space-time satisfying the weak energy condition, there exists a time-symmetric initial data set $(\bar{g}, 0)$ for an empty Einstein space-time (i.e., $\bar{R} = 0$) which has smaller or equal energy, if*

$$g - \hat{e} \in H^2_\delta, \qquad k \in H^2_{\delta-1},$$

for some $-\frac{3}{2} < \delta < -\frac{1}{2}$, $R(g)$ integrable.

Proof. set $\bar{g} = \phi^4 g$, with $\phi - 1 \in H^2_\delta$. Then also $\bar{g} - e \in H^2_\delta$ and we have, if $\Delta_g \phi$ is integrable,

$$\bar{E} = E - 4 \int_{\partial\Sigma} D^i \phi \, \mathrm{d}S_i = E - 4 \int_\Sigma \Delta_g \phi \, \mathrm{d}\Sigma \tag{2.3}$$

(the surface integral and $D^i\phi$ are computed equivalently with g or $\hat{e}$). The classical identity

$$\bar{R} \equiv 8\phi^{-5}(\phi R - \Delta_g \phi) \tag{2.4}$$

shows on the other hand that $\bar{R} = 0$ if

$$\Delta_g \phi - R\phi = 0, \tag{2.5}$$

i.e. $\Delta_g(\phi - 1) - R(\phi - 1) = R,$

which has one solution ϕ, with $\phi - 1 \in H^2_\delta$. Moreover, by the maximum principle, ϕ is > 0 on Σ, and $\Delta_g \phi$ is integrable by (2.5).

We deduce from (2.3) and (2.5)

$$\bar{E} = E - 4\int_\Sigma R\phi \, d\Sigma \leqslant E.$$

2.3. Existence of maximal slices

Let a slice Σ_f of the space–time $(M, {}^{(4)}g)$, $M = \Sigma \times \mathbb{R}$, be defined by an equation $s(t, x) \equiv t - f(x) = 0$, $x \in \Sigma$, $t \in \mathbb{R}$. The area of a domain $D_f \subset \Sigma_f$ is the functional $({}^{(4)}g, f) \to \mathscr{A}({}^{(4)}g, f) \in \mathbb{R}$ given by:

$$\mathscr{A}({}^{(4)}g, f) = \int_{D_f} dD_f,$$

where dD_f is the volume element induced by ${}^{(4)}g$ on D_f, that is, $dD_f = (\det g_f)^{1/2} d^3x$, g_f metric induced on Σ_f by ${}^{(4)}g$. We vary f, and thus D_f, but leaving the boundary ∂D_f fixed. The first variation of $\mathscr{A}$ is the linear operator on $h = \delta f$ given by (cf. Lichnerowicz 1944, Avez 1965)

$$\mathscr{A}'({}^{(4)}g, f)h \equiv \int_{D_f} hP({}^{(4)}g, f) dD_f,$$

where $P({}^{(4)}g, f)$ is the mean extrinsic curvature of Σ_f as a submanifold of $(M, {}^{(4)}g)$. It is a nonlinear, second-order, partial differential operator on f. Its vanishing expresses that Σ_f is (locally, under the above-explained circumstances) an extremum (in fact a maximum if Σ_f is space-like) of the area. The corresponding partial differential operator reads, in local coordinates x^α on $\Sigma \times \mathbb{R}$,

$$P({}^{(4)}g, f) \equiv -\mu_f^{1/2} \gamma_f^{\alpha\beta} \nabla_\alpha \partial_\beta s = 0,$$

where $s(t, x) = t - f(x)$,

$$-\mu_f = g^{\lambda\mu} \partial_\lambda s \partial_\mu s = g^{00} - 2g^{0i} \partial_i f + g^{ij} \partial_i f \partial_j f,$$

and

$$\gamma_f^{\alpha\beta} = g^{\alpha\beta} - n^\alpha{}_f n^\beta{}_f,$$

where n_α is the unit normal to Σ_f. In particular, γ_f^{ij} are the contravariant components of the metric g_f induced by ${}^{(4)}g$ on Σ_f. When Σ_f is space-like we have $\mu_f > 0$ and γ_f^{ij} positive definite. The operator P is then elliptic. Because the elliptic character depends on the solution itself, it is difficult to obtain uniform *a priori* bounds on the solution, which, in usual elliptic theory, lead to existence theorems. In fact, for a general space-time, it is still not known* if it admits maximal slices. However, there is a stability property related to these slices, in the case of a maximal space-like submanifold of a hyperbolic manifold, which leads to the following theorem (Choquet-Bruhat 1976, cf. also Choquet-Bruhat et al. 1979), which we enunciate for asymptotically Minkowskian manifolds (though it is also valid in other contexts).

Theorem. *Let* $(M, {}^{(4)}\tilde{g})$ *be a space-time satisfying the strong energy condition, asymptotically Minkowskian in the sense that* $M = \Sigma \times \mathbb{R}$, $(\Sigma, \hat{e})$ *is Euclidean at infinity, and* ${}^{(4)}\tilde{g} - \hat{\eta} \in E^2_\delta$ (cf. § 1.2) *for some* δ *with* $-\frac{3}{2} < \delta < -\frac{1}{2}$, *which admits a maximal slice* Σ_0. *Then there exists a number* $c > 0$ *such that each neighboring space-time* ${}^{(4)}g$, *in the sense that*

$$\|{}^{(4)}g - {}^{(4)}\tilde{g}\|_{E^2_\delta} < c,$$

admits a maximal slice.

Proof. The mapping $({}^{(4)}g, f) \mapsto P({}^{(4)}g, f)$ is a C^∞ mapping into $E^0_{\delta+2}$ if ${}^{(4)}g - \eta \in E^2_\delta$ and $f \in H^2_\delta$. Its derivative with respect to f at the point ${}^{(4)}g = {}^{(4)}\tilde{g}$, $f = 0$ (we take the original slicing $\Sigma \times \mathbb{R}$ to be $\tilde{g}$-orthogonal at Σ_0, that is, $\delta f = h$ to be taken in a direction orthogonal to Σ_0) is computed to be

$$P'_f({}^{(4)}\tilde{g}, 0) \cdot h \equiv \tilde{\Delta}_0(\alpha h) - \alpha h(\tilde{k}_0 \cdot \tilde{k}_0 + \tilde{R}_{\lambda\mu}\tilde{n}_0^\lambda \tilde{n}_0^\mu).$$

$\tilde{\Delta}_0 = \Delta_{\tilde{g}_0}$ is the Laplace operator of the metric $\tilde{g}_0$ induced on Σ_0 by ${}^{(4)}\tilde{g}$. If the metric ${}^{(4)}\tilde{g}$ satisfies the strong energy condition the operator $P'_f({}^{(4)}\tilde{g}, 0)$ is of the form $\Delta_g + f$ with $f \leqslant 0$. It is then easy to deduce from the theorem in § 2.1 that, given ${}^{(4)}g \in E^2_\delta$, this operator is an isomorphism $H^2_\delta \to E^0_{\delta+2}$. By the implicit function theorem the conclusion follows.

Corollary. *There is a neighborhood (in the sense of the theorem) of*

* Recent proof for asymptotically Minkowskian space-times: Bartnik (1984).

Minkowski space-time such that all space-times in this neighborhood admit a maximal slice.

2.4. The mass functional and its derivatives

By the O'Murchadha and York theorem (§2.2) it is sufficient, for space-times in a neighborhood of Minkowski space-time, to prove the positivity of energy for time-symmetric initial data $(g, 0)$, where g has vanishing scalar curvature,

$$R(g) = 0. \tag{2.6}$$

For time-symmetric initial data the momentum P^k is zero and the mass is identical with the energy.

We first prove

Theorem. *The subset of three-metrics on $(\Sigma, \hat{e})$ such that $R(g) = 0$, $g - \hat{e} \in H^2_\delta$, $-\frac{3}{2} < \delta < -\frac{1}{2}$ forms a C^∞ submanifold $\mathscr{C}$ of H^2_δ.*

Proof. We do not have here to suppose that Σ is diffeomorphic to $\mathbb{R}^3$. $g \mapsto R(g)$ is a C^∞ map (cf. multiplication theorem), $g - \hat{e} \in H^2_\delta \mapsto R(g) \in H^0_{\delta+2}$. Its derivative (first variation) at a point g is the linear map acting on $h = \delta g$,

$$R'(g) \cdot h \equiv -\Delta_g \operatorname{tr} h + D_i D_j h^{ij} - R_{ij} h^{ij}; \tag{2.7}$$

this mapping is surjective at any point g such that $R(g) = 0$. Let $f \in H^0_{\delta+2}$, set

$$h = \tfrac{1}{3} g \tau \tag{2.8}$$

with τ a scalar function. Then $R'(g) \cdot h = f$ reads

$$-\tfrac{2}{3} \Delta_g \tau = f \tag{2.9}$$

which has a solution $\tau \in H^2_\delta$ by the theorem in §2.1. The uniqueness and continuous dependence on f of this solution shows that every $h \in H^2_\delta$ admits a unique continuous decomposition of the form

$$h = h_1 + h_2$$

with

$$h_1 \in \ker R'(g), \quad \text{that is,} \quad R'(g) h_1 = 0,$$

and h_2 of the form (2.8).

The C^∞ mapping $g \mapsto R(g)$ is a submersion at points where $R(g) = 0$ which is therefore a C^∞ submanifold (cf. for instance Choquet-Bruhat et al. 1982).

Remark. The fact that the solution of the constraints, not necessarily time-symmetric but with $\operatorname{tr} k = 0$, forms a submanifold in the asymptotically Euclidean case is also true (Choquet-Bruhat et al. 1977), due to the absence of asymptotically vanishing Killing fields on such manifolds (Choquet-Bruhat and Choquet 1979, another proof in Christodoulou and O'Murchadha 1981).

We shall define the energy functional $E(g)$ for the metric g, with $R(g) = 0$ as the restriction to $\mathscr{C}$ of a C^∞ functional on a Hilbert space H^2_δ, namely H^2_{-1}. This functional is

$$g - \hat{e} \in H^2_{-1} \mapsto \tilde{E}(g) \in R \quad \text{with}$$

$$\tilde{E}(g) = \int_\Sigma (\partial^i \partial^j g_{ij} - \partial^i \partial^i g_{jj}) \mathrm{d}\Sigma(\hat{e}) - \int_\Sigma R(g) \mathrm{d}\Sigma(g). \tag{2.10}$$

(Derivatives ∂ and raising of indices done with $\hat{e}$ in the first integral. Note that only the difference of the two integrals is defined on H^2_{-1}). The derivative of this functional at an arbitrary point g, with $g - \hat{e} \in H^2_{-1}$, is

$$\tilde{E}'(g) \cdot h = \int_\Sigma (R_{ij} - \tfrac{1}{2} g_{ij} R) h^{ij} \mathrm{d}\Sigma(g); \tag{2.11}$$

thus, if $g \in \mathscr{C}$;

$$\tilde{E}'(g) \cdot h = \int_\Sigma R_{ij} h^{ij} \mathrm{d}\Sigma(g). \tag{2.12}$$

A critical point of E is a metric g such that (2.12) vanishes for every $h \in \mathrm{T}_g \mathscr{C}$, the tangent space to $\mathscr{C}$ at g, the space of those h which satisfy

$$R'(g)h \equiv -\Delta_g \operatorname{tr} h + D_i D_j h^{ij} - R_{ij} h^{ij} = 0. \tag{2.13}$$

As we have just seen, from an arbitrary $h \in H^2_\delta$ we deduce an $h_1 \in \mathrm{T}_g \mathscr{C}$ by taking

$$h_1 = h + \tfrac{1}{3} g \tau$$

with

$$\Delta_g \tau = \tfrac{3}{2}(-\Delta_g \operatorname{tr} h + D_i D_j h^{ij} - R_{ij} h^{ij}).$$

At a point $g \in \mathscr{G}$ the condition $\tilde{E}'(g) \cdot h_1 = 0, \forall h_1 \in T_g\mathscr{G}$, is therefore equivalent to

$$\int_\Sigma R_{ij} h^{ij} d\Sigma(g) = 0 \quad \forall h \in H^2_{-1}, \tag{2.14}$$

and therefore $R_{ij} = 0$.

We conclude that the flat metrics are the only critical points of the energy functional on $\mathscr{G}$. If $\Sigma = \mathbb{R}^3$, the critical points are the orbit of the Euclidean metric under diffeomorphisms (asymptotic to the identity). More generally we denote by $\hat{e}$ a chosen flat metric on Σ (supposed to exist) Euclidean at infinity—for simplicity we identify it with the metric used in the definitions of H^s_δ spaces on Σ (by results of differential geometry, if $\hat{e}$ exists, Σ is $\mathbb{R}^3$).

A straightforward computation shows that the second derivative of $\tilde{E}$ at an arbitrary point g is the quadratic form on h

$$\begin{aligned}\tilde{E}''(g)(h,h) = &\int_\Sigma \tfrac{1}{2}(Dh \cdot Dh) d\Sigma(g) \\ &- \int_\Sigma \{\tfrac{1}{2}(D \operatorname{tr} h)^2 + (D \cdot h)^2 + D \operatorname{tr} h \cdot (D \cdot h)\} d\Sigma(g) \\ &+ \int_\Sigma \{-\operatorname{tr} h(R_{ij} h^{ij}) + R_{ij} h^{ik} h^j{}_k + \tfrac{1}{4} R (\operatorname{tr} h)^2\} d\Sigma(g).\end{aligned} \tag{2.15}$$

At flat space $R_{ij} = 0$. If moreover we restrict h by the conditions (traceless, transverse h)

$$\operatorname{tr} h = 0, \qquad D \cdot h = 0, \tag{2.16}$$

the second derivative of $\tilde{E}$ reduces to a positive quadratic form (f denotes a flat metric):

$$\tilde{E}''(f)(h,h) = \tfrac{1}{2} \int_\Sigma Dh \cdot Dh \, d\Sigma(g). \tag{2.17}$$

We remark that, if $h \in T_f\mathscr{G}$, it satisfies

$$-\Delta_f \operatorname{tr} h + D_f \cdot D_f \cdot h = 0. \tag{2.18}$$

If we impose on h the condition

$$D_f \cdot h - \tfrac{1}{2}\mathrm{d}\,\mathrm{tr}_f h = 0, \tag{2.19}$$

it will also satisfy, if in $T_f\mathscr{C}$, the transverse–traceless conditions (2.16). The condition (2.19) is easily seen to be the principal part (for $g = f$) of the variation of the following restriction on g:

$$H(g) \equiv g \cdot (\Gamma(g) - \Gamma(\hat{e})) = 0 \quad [\text{i.e., } g^{ij}(\Gamma^k_{ij} - \Gamma^k_{ij}(\hat{e})], \tag{2.20}$$

which expresses that the identity mapping on Σ is a harmonic map $(\Sigma, g) \to (\Sigma, \hat{e})$. In fact the first variation of (2.20) is

$$H'(g) \cdot h \equiv D \cdot h - \tfrac{1}{2}\mathrm{d}(\mathrm{tr}\, h) - h \cdot S = 0,$$

where S is the three-tensor difference of the connection of g and $\hat{e}$.

It can be proved that, at least in a neighborhood of $\hat{e}$, the subset $H(g) = 0$ of $\mathscr{C}$ is again a C^∞ submanifold, called $\mathscr{S}$ of H^2_δ.

The result we have is thus that the energy functional E on $\mathscr{S}$ has a critical point at $\hat{e}$ and that its second derivative is a positive quadratic form in $T_{\hat{e}}\mathscr{S}$: it vanishes only, if $h \in H^2_{-1}$ and $h \in T_{\hat{e}}\mathscr{S}$, when $h = 0$. However, it is not positive definite in the H^2_{-1} sense, since the second derivatives of h do not appear.

However, by examination of the general formula (2.15) and Sobolev-type estimates, it is possible to show that the second derivative of the energy for metrics $g \in \mathscr{S}$ is positive, at least in a neighborhood of $\hat{e}$. From this fact the positivity of E in this neighborhood can be deduced with the help of a Taylor formula. We even obtain an inequality of the type (with C some Sobolev constant)

$$E(g) \geqslant C\|g - \hat{e}\|^2 = C \int_\Sigma |\partial(g - \hat{e})|^2 \mathrm{d}\Sigma(e)$$

($\|\ \|$ denotes the "energy norm").

It can also be proved that every metric g, $g - \hat{e} \in H^3_\delta$ sufficiently near to $\hat{e}$ is isometric to a metric satisfying the harmonicity condition, thus also satisfies $m(g) \geqslant 0$, and $m(g) = 0$ only if g is isometric to $\hat{e}$.

One of the difficulties to extend this proof of the positivity of the mass to all asymptotically flat spaces rests on a fact familiar in global analysis, that in infinite-dimensional vector spaces all norms are not equivalent. Another difficulty, of geometric origin, is that $\mathscr{C}$ apparently does not admit a global slice for the group of isometries (corresponding to the Gribov

ambiguity in Yang–Mills theory). Other ways to factor out the isometry group could perhaps be used (cf. the work of Atiyah and Bott 1983 on the critical points of the Yang–Mills functional).

3. The proof of Schoen and Yau

The basic theorem of Schoen and Yau (1979) about the positive energy of asymptotically Minkowskian space–times is purely geometrical: it says that an asymptotically Euclidean three-manifold with metric of the type

$$g_{ij} = \left(1 + \frac{m}{2r}\right)^4 \delta_{ij} + h_{ij}, \quad h_{ij} = O(r^{-2}) \tag{3.1}$$

cannot have both $R(g) \geqslant 0$ and $m < 0$. It thus applies to space–times which satisfy the weak energy condition and admit a maximal slice ($\operatorname{tr} k = 0$). They relieved this last condition in a later paper (Schoen and Yau 1981b), which contains also deep ideas concerning horizons.

3.1. Second variation of the area

The proof of the basic theorem relies on the consideration of minimal (with respect to area) two-dimensional surfaces of the three-space (Σ, g). Indeed, the existence of such a minimal surface S yields, by considering the second variation of the area $\mathscr{A}$, a functional inequality

$$\mathscr{A}''(S) \geqslant 0. \tag{3.2}$$

The analogue of the inequality (3.2) for one-dimensional submanifolds which minimize (maximize in the case of a hyperbolic metric) the arc length has been the source of many important global properties in Riemannian geometry—and is the cornerstone of the singularity theorems in General Relativity. Like the second variation around a geodesic, the second variation around a minimal (or maximal) submanifold has a remarkably simple expression in geometrical terms (we have seen such an expression in § 2.3). However, the existence of these submanifolds, with the relevant properties, is difficult to prove, and that is why the Schoen and Yau proof, though conceptually simple, is technically very intricate. I shall sketch the main ideas, and hint at some of the technicalities.

The second variation of the area of S, $\mathscr{A}''(S)$, is a quadratic form on

functions of class C^2 on S which for functions with compact support reads

$$\mathscr{A}''(S)(f,f) = -\int_S f(\Delta_S f + f\,\mathrm{Ricc}(v,v) + f|b|^2)\mathrm{d}S$$

$$= \int_S (|\nabla f|^2 - f^2(\mathrm{Ricc}(v,v) + |b|^2)\mathrm{d}S, \tag{3.3}$$

where the Laplacian Δ_S, the norm $|\ |$ and $\mathrm{d}S$ are taken in the metric induced on S by g, $\mathrm{Ricc}(v,v)$ is the Ricci tensor of g contracted with the unit normal v to S, and b is the extrinsic curvature of S in (Σ, g). One has the identity (cf. an analogous one in the Hamiltonian constraint), if $\mathrm{tr}\, b = 0$ (verified since S is a minimal surface),

$$\mathrm{Ricc}(v,v) - R = -\tfrac{1}{2}|b|^2 - R_S, \tag{3.4}$$

where R_S is the scalar (Gaussian) curvature of S; thus (3.3) can be written

$$\int_S (R + \tfrac{1}{2}|b|^2 - R_S) f^2 \mathrm{d}S. \tag{3.5}$$

3.2. Metrics Euclidean outside of a compact set

If S is compact we can take $f = 1$. For such an f the condition $\mathscr{A}'' \geqslant 0$ gives, using the identity (3.5) and the Gauss–Bonnet theorem,

$$\int_S (R + \tfrac{1}{2}|b|^2)\mathrm{d}S \leqslant \int_S R_S \mathrm{d}S = 2\pi\chi(S), \tag{3.6}$$

where $\chi(S)$ is the Euler–Poincaré characteristic. The inequality (3.6) gives a relation between the topological type of S and the scalar curvature of Σ. We have

Theorem. *If (Σ, g) is a compact manifold which admits a minimal submanifold S with $\chi(S) = 0$, and the conformal manifolds $(\Sigma, \phi^4 g)$ admit also a minimal submanifold S_ϕ with $\chi(S_\phi) = 0$, then if $R(g) \geqslant 0$, g is flat.*

Proof. The inequality (3.6) is incompatible with $R > 0$. But if $R \geqslant 0$ and $\mathrm{Ricc}(g) \neq 0$ there exists a conformal manifold (Σ, g_1) with $R(g_1) > 0$ (cf. Kazdan and Warner 1975).

Thus $\mathrm{Ricc}(g) = 0$ (and $R = 0$).

Corollary. *On a torus* T^3 *there exists no non-flat metric* g *with* $R \geqslant 0$.

Proof. Each Riemaniann torus (T^3, g) admits a minimal submanifold diffeomorphic to T^2, and a torus has vanishing Euler–Poincaré characteristic.

The previous theorem leads to a very simple proof of the following:

Theorem. *If* g *is a metric on* $\mathbb{R}^3$, *Euclidean outside of a compact set, with* $R(g) \geqslant 0$, *then* g *is flat* (*and* $R(g) \equiv 0$).

Proof. Place the compact in a cube, extend the cube periodically, and by identification construct a torus T^3 with a smooth metric.

Remark. The theorem gives a class of metrics, with $m = 0$, which must be flat if $R(g) \geqslant 0$, i.e. for maximal space-like sections of an Einsteinian space–time satisfying the weak energy condition.

3.3. Asymptotically Euclidean manifolds

The basic theorem of Schoen and Yau (1979), relative to asymptotically Euclidean manifolds, is

Theorem. *Let* (Σ, g) *be a* (*oriented, three-dimensional*) *Riemaniann manifold, with* $R(g) \geqslant 0$; *asymptotically Euclidean in the sense that in each end* $\Omega_\alpha \simeq \mathbb{R}^3 \setminus \bar{B}_{\rho_\alpha}$ *the metric is, in the Euclidean coordinates of* $\mathbb{R}^3$,

$$g_{ij} = \left(1 + \frac{m_\alpha}{2r}\right)^4 \delta_{ij} + h_{ij}. \tag{3.7}$$

(A) *If* $h \in C^2_2$ (*cf. definitions* § 2), then each m_α *is non-negative.*

(B) *If furthermore* $h \in C^5$ *and* $\sigma^5(|\partial^3 h| + |\partial^4 h| + |\partial^5 h| < +\infty$ *and one* m_α *is zero, then* (Σ, g) *is isometric to* $\mathbb{R}^3$ *with the flat metric.*

Proof. (A) Schoen and Yau suppose that in one end, say Ω_0, the mass m_0, denoted simply m, is negative. Thier proof then goes in three steps.

(1) They use the assumption $m < 0$ to prove the existence of a complete area-minimizing surface S, properly imbedded in Σ so that $S \cap \{\Sigma \setminus \Omega_0\}$ is compact and $S \cap \Omega_0$ lies between two parallel Euclidean two-planes, when considered in the coordinates (x^i) of $\mathbb{R}^3$. These properties of S will

be needed to show in step (3) the contradiction between the existence of S, the second variation argument, and the other hypotheses.

The surface S is constructed as a limit of two-surfaces S_ρ, $\rho \rightsquigarrow \infty$, with each S_ρ of least area among the oriented surfaces with boundary curve the Euclidean circle C_ρ of center 0 in the plane $x^3 = 0$ of $\mathbb{R}^3$, and radius ρ large enough for C_ρ to lie in Ω_0. Schoen and Yau show that the S_ρ cannot run to infinity in an other end than Ω_0, if these other ends are also asymptotically Euclidean. They also show that if the mass m relative to Ω_0 is negative [$m_0 < 0$ in (3.4)] then all the $S_\rho \cap \Omega_0$ lie between two parallel planes $x^3 = \pm\bar{h}$. These two properties are proved through the maximum principle. This makes use of the fact that, at a point of $S_\rho \cap \Omega_0$, where $x^3 = \bar{h}$ is extremum one has, using the fact that S_ρ is minimal, and denoting by q^{ij} the metric induced by g on S:

$$q^{ij}D_iD_jx^3 = -\frac{2m\bar{h}}{r^2} + O(r^{-3}), \tag{3.8}$$

which is incompatible with $\bar{h}$ too large, $m < 0$ and $\bar{h}$ a maximum of x^3.

(2) If g is asymptotically flat on Σ with $R \geqslant 0$ and the mass m of g relative to Ω_0 is negative, then there is $\tilde{g} = \phi^4 g$ having $\tilde{R} \geqslant 0$, $\tilde{R} > 0$ outside a compact subset of Ω_0, and such that $\hat{g}$ has negative mass relative to Ω_0.

The proof is by considering, in Ω_0, the g-Laplacian of r^{-1}; it reads, using the condition (3.7) on g,

$$\begin{aligned}\Delta_g \frac{1}{r} &= \sum_{i=1}^{3} \frac{\partial}{\partial x^i}\left(1+\frac{m}{2r}\right)^2 \frac{\partial}{\partial x^i}\frac{1}{r} + O\left(\frac{1}{r^5}\right) \\ &= \frac{m}{r^4} + O\left(\frac{1}{r^5}\right).\end{aligned}$$

It follows that there is a number $\rho > \rho_0$ so that

$$\Delta_g \frac{1}{r} < 0 \quad \text{for } r \geqslant \rho. \tag{3.9}$$

Schoen and Yau then define a smooth positive function $\phi > 0$ on Σ, constant on $\Sigma \setminus \Omega_0$, such that in Ω_0, for large enough r (for instance $r > 2\rho_0$)

$$\phi = 1 - \frac{m}{4r},$$

and which satisfies $\Delta\phi \leqslant 0$ on Σ and, as a consequence of (3.9) and

$m < 0$,

$$\Delta\phi < 0 \quad \text{in } \Omega_0 \cap \{r > 2\rho_0\}. \tag{3.10}$$

In Ω_0, for r large enough ($r > 2\rho_0$), we have

$$\tilde{g}_{ij} = \left(1 + \frac{m}{4r}\right)^4 \delta_{ij} + O(r^{-2});$$

thus the mass $\tilde{m}$ of $\tilde{g}$ is

$$\tilde{m} = \tfrac{1}{2}m < 0,$$

while $\tilde{R}$ is

$$\tilde{R} = 8\phi^{-5}(R - \Delta_g\phi) \geqslant 0,$$

and thus, by (3.10),

$$\tilde{R} > 0 \quad \text{in } \Omega_0 \cap \{r > 2\rho_0\}.$$

(3) Schoen and Yau use a cut-off function f, equal to 1 on

$$S_{(\rho)} = \{S \cap (\Sigma \setminus \Omega_0)\} \cup \{S \cap B_\rho\}$$

and vanishing outside of $S_{(\rho^2)}$, inserted in (3.5) to show, after making appropriate estimates and taking limits, that

$$\int_S R_S \mathrm{d}S > 0$$

(recall that for $\tilde{g}$, now denoted g, we have $R > 0$ outside a compact subset of S). They give then two different proofs that in the situation at hand one has

$$\int_S R_S \mathrm{d}S \leqslant 0.$$

These proofs largely use the Gauss–Bonnet theorem with boundary, estimates of lengths on this boundary, and some deep results of differential geometry.

The proof that $m = 0$ for Ω_0 implies g flat involves the construction in case $R \geqslant 0$ and $R \not\equiv 0$, of a conformal metric $\hat{g}$, with $\tilde{R} = 0$ and $\tilde{m} < 0$ (theorem of O'Murchadha and York, cf. §2): therefore part (A) of the theorem shows that $m = 0$ implies $R \equiv 0$. Schoen and Yau suppose then $\mathrm{Ricc}(g) \not\equiv 0$, and show that it is then possible to construct a one-

parameter family of metrics with $R(g_t) = 0$ and such that

$$g_t = \phi_t^4(g + t\,\mathrm{Ricc}(g)).$$

Their energy is

$$m(t) = -\frac{1}{32\pi}\int_\Sigma R_t\phi_t\,\mathrm{d}\mu(g_t), \quad m(0) = m = 0,$$

and, by using the various variation formulas for R and $\mathrm{Ricc}(g)$,

$$\left.\frac{\mathrm{d}m}{\mathrm{d}t}\right|_{t=0} = \frac{1}{32\pi}\int_\Sigma |\mathrm{Ricc}(g)|^2\,\mathrm{d}\mu(g);$$

thus $m'(0) > 0$ if $\mathrm{Ricc}(g) \neq 0$, which would imply, for a suitable $t < 0$, $m(t) < 0$, in contradiction with the first part of the theorem.

Remark. The asymptotic behavior (3.1) of the metric g is more restrictive than the usual one (cf. § 2), since it supposes that the $O(r^{-1})$ deviation of g from the Euclidean metric is spherically symmetric, of the form $2mr^{-1}$, like the Schwarzschild metric in isotropic coordinates. Such an asymptotic behavior had been adopted by Geroch (1973) as a starting point for the very simple method he proposed to prove the positive-mass conjecture, using a family of nested topological two-spheres and the Gauss–Bonnet formula. The existence of these two-spheres, satisfying the equation required by Geroch and the innermost reducing to a point, is however dubious except in special cases (for instance spherical symmetry or flat g, "pure kinetic case", Pong 1976).

It has been pointed out by York (1980) that the definition (3.1) does not include all cases of phsysical interest; for example, the metric of a boosted slice of a Schwarzschild space–time is not of this type (note that the coefficient m in (3.1) is in fact the energy as defined in § 2). In a subsequent paper Schoen and Yau (1981a) showed that their result applies to the more general asymptotic behavior, by a perturbation argument.

3.3.1. Extension of the Schoen–Yau proof to the case $\mathrm{tr}\,k \neq 0$

In a second paper Schoen and Yau (1981b) prove the positivity of the energy for an asymptotically Euclidean initial set of a space–time which verifies the dominant energy condition, without hypothesis on the

existence of a maximal slice [they only suppose that $\mathrm{tr}\, k$ has fall-off $O(r^{-3})$, which is physically reasonable by York's 1980 analysis].

Their idea is to define the initial data set (Σ, g, k) in two steps. The first step is to find a hypersurface $\tilde{\Sigma}$ in $\Sigma \times \mathbb{R}$, $t = w(x)$, which satisfies the equation proposed by Pong (1978),

$$\left(k^{ij} - \frac{D^i D^j W}{1+|DW|^2}\right)\left(g_{ij} - \frac{D_i W D_j W}{1+|DW|^2}\right) = 0, \tag{3.11}$$

and is asymptotic to Σ in such a way that it has the same energy. A motivation for considering such an equation is that if Σ is a space-like slice of Minkowski space-time, $\tilde{\Sigma}$ can be identified with a maximal slice of that space-time.

The second step is to show that if such a slice $\tilde{\Sigma}$ exists the induced metric on $\tilde{\Sigma}$ can be conformally deformed to one with zero scalar curvature and smaller or equal energy, to which their basic theorem applies.

It happens that the slice $\tilde{\Sigma}$ does not exist in general: it is remarkable that its existence is related to the existence of apparent horizons in the initial data set (even if it is assumed to be non-singular): Schoen and Yau prove the existence of a solution of (3.11), asymptotic to zero at infinity, defined outside a finite family of apparent horizons: a boundary S of such a horizon satisfies the classical condition

$$K_S - \mathrm{tr}_S k = 0 \qquad \text{or} \quad K_S + \mathrm{tr}_S k = 0,$$

where K_S is the mean curvature of S in Σ and $\mathrm{tr}_S k$ the trace of the restriction of k to S.

They study the asymptotic behavior on Σ of the apparent horizons, and are thus able to perform the second step of their proof.

As their proof shows, the positive-energy theorem of Schoen and Yau applies to initial data with singularities if these are hidden inside apparent horizons.

4. Witten's proof

The ideas in this proof are so elegant and simple that it is possible to give it in a completely rigorous and transparent way. In the course of the exposition we will also use complements and clarification due to Parker and Taubes (1982), Nester (1981) and Reula (1982). The existence of a solution of "Witten's equation" is also a consequence of the general theorem proved by Choquet-Bruhat and Christodoulou (1981) for elliptic systems on an asymptotically Euclidean manifold.

4.1. Definition of the Witten two-form

We first give a two-form ω, defined on space–time, whose integral over the "two-sphere at infinity" S in an end of a space-like slice of an asymptotically Minkowskian space–time is $P_\lambda U^\lambda$, with P_λ the energy–momentum (cf. § 1) and U^λ a time-like vector. It is

$$\omega = \tfrac{1}{2} E^{\sigma\alpha} \mathrm{d}S_{\sigma\alpha}, \tag{4.1}$$

with $\mathrm{d}S_{\sigma\alpha}$ the two-dimensional volume element corresponding to the metric ${}^{(4)}g$,

$$\mathrm{d}S_{\sigma\alpha} = \tfrac{1}{2}\eta_{\sigma\alpha\lambda\mu}\, \mathrm{d}x^\lambda \wedge \mathrm{d}x^\mu, \tag{4.2}$$

and $E^{\sigma\alpha}$ the real antisymmetric tensor:

$$E^{\sigma\alpha} = 2\eta^{\sigma\alpha\delta\beta}(\bar{\psi}\xi\gamma_\delta \nabla_\beta \psi - \nabla_\beta \bar{\psi}\xi\gamma_\delta \psi); \tag{4.3}$$

ψ is a Dirac four-spinor, the Dirac matrices are standard in an orthonormal frame, $\gamma^a\gamma^b + \gamma^b\gamma^a = 2\eta^{ab} I$, γ^0 antihermitian and γ^i Hermitian, $\bar{\psi} = \bar{\psi}\gamma^0$ the Dirac adjoint covariant spinor, $\xi = (1/4!)\eta_{\alpha\beta\lambda\mu}\gamma^\alpha\gamma^\beta\gamma^\lambda\gamma^\mu$.

ω can also be written as

$$\omega = 4(\bar{\psi}\xi\gamma_\lambda \nabla_\mu \psi - \nabla_\mu \bar{\psi}\xi\gamma_\lambda \psi)\mathrm{d}x^\lambda \wedge \mathrm{d}x^\mu. \tag{4.4}$$

Lemma 1. *If on Σ we have*

$$\psi = \psi_0 + \psi_1 \tag{4.5}$$

with ψ_0 smooth and constant on $\Sigma \setminus K$ (i.e., $\partial\psi_0 = 0$) and vanishing in other ends than Ω of Σ, while $\psi_1 \in H^1_\delta$, $\delta < -\frac{1}{2}$ (for instance $\delta = -1$), then

$$\int_S \omega = P_\lambda U^\lambda \quad \textit{with} \quad U^\lambda = \bar{\psi}_0 \gamma^\lambda \psi_0. \tag{4.6}$$

Proof. We have $\omega = \omega_0 + \omega_1$ with, on Σ ($x^0 = 0$),

$$\omega_0 = 4(\bar{\psi}_0 \xi\gamma_i \nabla_j \psi_0 - \nabla_j \bar{\psi}_0 \xi\gamma_i \psi_0)\mathrm{d}x^i \wedge \mathrm{d}x^j, \tag{4.7}$$

$$\omega_1 = 4(\bar{\psi}_0 \xi\gamma_i \nabla_j \psi_1 - \nabla_j \bar{\psi}_1 \xi\gamma_i \psi_0 + \bar{\psi}_1 \xi\gamma_i \nabla_j \psi - \bar{\nabla}_j \psi \xi\gamma_i \psi_1)\mathrm{d}x^i \wedge \mathrm{d}x^j. \tag{4.8}$$

We prove first that

$$\int_{\partial\Sigma} \omega_1 = 0. \tag{4.9}$$

Indeed, by definition (K_n sequence of increasing compact sets covering Σ),

$$\int_{\partial\Sigma} \omega_1 = \lim_{n=\infty} \int_{\partial K_n} \omega_1 = \lim_{n=\infty} \int_{K_n} d\omega_1 = \int_{\Sigma} d\omega_1 ; \tag{4.10}$$

we have, on Σ,

$$d\omega_1 = 8\,\mathrm{Re}\{\nabla_k\bar{\psi}_0\xi\gamma_i\nabla_j\psi_1 + \bar{\psi}_0\xi\gamma_i\nabla_k\nabla_j\psi_1 + \nabla_k\bar{\psi}_1\xi\gamma_i\nabla_j\psi + \bar{\psi}_1\xi\gamma_i\nabla_k\nabla_j\psi\}dx^k \wedge dx^i \wedge dx^j. \tag{4.11}$$

The second covariant derivatives, antisymmetrized, are replaced by the Riemann tensor whose restriction to Σ is in H^0_1. An inspection of the various terms added in (4.11) shows that they are all integrable on Σ, and that, given $\bar{\psi}_0$, their integral on Σ depends continuously on the H^1_{-1} norm of ψ_1. We can therefore obtain $\int_\Sigma d\omega_1$ as the limit of integrals of exact differentials of forms with compact support (by approximating ψ_1 in H^1_{-1} by spinors with compact support): such integrals are zero on Σ, and the same is true of their limit.

Second, we prove that

$$\int_S \omega_0 = P_\lambda U^\lambda, \tag{4.12}$$

where U^λ is the time-like or null vector past-directed in the Lorentzian frame at infinity, given by

$$U^\lambda = \bar{\psi}_0\gamma^\lambda\psi_0. \tag{4.13}$$

Indeed, by hypothesis ψ_0 is constant on $\Sigma \setminus \Omega$, that is, $\partial_i\psi_0 = 0$. Thus, on $\Sigma \setminus \Omega$:

$$\omega_0 = -\bar{\psi}_0(\xi\gamma_i\sigma_{ab} + \sigma_{ab}\xi\gamma_i)\psi_0\omega_j{}^{ab}dx^i \wedge dx^j,$$

which gives, by Dirac algebra,

$$\omega_0 = \bar{\psi}_0(2\eta_{iab\lambda}\gamma^\lambda)\psi_0\omega_j{}^{ab}dx^i \wedge dx^j = 2u^\lambda\pi_\lambda, \tag{4.14}$$

where π_λ is the covector-valued two-form

$$\pi_\lambda = \eta_{iab\lambda}\omega_j{}^{ab}dx^i \wedge dx^j ; \tag{4.15}$$

we show that the integrals on S of π^0 and π^k give, respectively, the

expressions (1.24) and (1.25) for P^0 and P^k. Indeed (on S we have $\eta_{ihk}{}^0 = -\eta_{ihk}$),

$$\int_S \pi^0 = \int_S \eta_{ihk}{}^0 \omega_j{}^{hk} dx^i \wedge dx^j = -\int_S \eta^{ijl}\eta_{ihk}\omega_j{}^{hk} dS_l$$

$$= -\int_S \varepsilon^{jl}_{hk}\omega_j{}^{hk} ds_l = 2\int_S \omega_j{}^{lj} ds_l. \tag{4.15a}$$

If the metric is, in $\Sigma \setminus K$,

$$g_{\lambda\nu} = \eta_{\lambda\mu} + h_{\lambda\mu}, \quad h_{\lambda\mu} \in H^2_\delta, \quad \partial_\alpha h_{\lambda\mu} \in H^1_{\delta+1},$$

a moving Lorentzian coframe in $\Sigma \setminus K$, $g_{\lambda\mu}dx^\lambda dx^\mu \equiv \eta_{\alpha\beta}\theta^\alpha\theta^\beta$, is $\theta^\alpha = dx^\alpha + a^\alpha{}_\beta dx^\beta$ with

$$a^\alpha{}_\beta = \tfrac{1}{2}\eta^{\alpha\lambda}(h_{\lambda\beta} + f_{\lambda\beta}), \quad f_{\lambda\beta} \in H^2_{\delta+1}, \quad \partial_\alpha f_{\lambda\mu} \in H^1_{\delta+2}.$$

The standard expression for the Riemannian connection in an orthonormal frame (cf. for instance Choquet-Bruhat et al. 1982, p. 308) then gives (terms coming from $f_{\alpha\beta}$ do not contribute to the integral on S)

$$\int_S \pi^0 = \int_S (\partial_j h_{ij} - \partial_i h_{jj}) dS_i = P^0$$

and

$$\int_S \pi_k = \int_S \eta_{iabk}\eta^{ijl}\omega_j{}^{ab} dS = -2\int_S \eta_{ihk}\eta^{ijl}\omega_j{}^{0h} ds_l$$

$$= -2\int_S (\omega_j{}^{0j}\delta_k{}^l - \omega_k{}^{0l}) dS_l = -2\int_S (\delta^i_k \operatorname{tr} k - k^i_k) dS_i$$

$$= P^k.$$

4.2. Proof of positivity

From the Stokes theorem and the expression (4.1) for ω we deduce*:

* $d\omega$ will be integrable on Σ, for metrics such that their source $T_{\lambda\mu}$ is integrable, but ψ_0 cannot be approximated by functions with compact support in a norm such that the integral of $d\omega$ depends continuously on ψ_0 in this norm, thus $\int_{\partial\Sigma}\omega$ is not necessarily zero!

$$\int_{\partial\Sigma}\omega=\int_{\Sigma}\mathrm{d}\omega=\int_{\Sigma}\nabla_\alpha E^{\sigma\alpha}\mathrm{d}\Sigma_\sigma,\qquad \mathrm{d}\Sigma_\sigma=\frac{1}{3!}\eta_{\sigma\beta\lambda\mu}\,\mathrm{d}x^\beta\wedge\mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu \tag{4.16}$$

with (due to the antisymmetry in ρ, β this expression is real)

$$\nabla_\rho E^{\sigma\rho}=8\eta^{\sigma\rho\delta\beta}(\nabla_\rho\bar\psi\xi\gamma_\delta\nabla_\beta\psi+\bar\psi\xi\gamma_\delta\nabla_\rho\nabla_\beta\psi).$$

We transform the first term by the Dirac algebra property

$$\eta^{\sigma\alpha\delta\beta}\xi\gamma_\delta=\tfrac{1}{2}(\gamma^\sigma\gamma^{\alpha\beta}+\gamma^{\alpha\beta}\gamma^\sigma),$$

and the second by the Ricci identity on spinors (cf. Lichnerowicz 1964)

$$\nabla_\alpha\nabla_\beta\psi-\nabla_\beta\nabla_\alpha\psi=-\tfrac{1}{4}R_{\alpha\beta}{}^{\lambda\mu}\gamma_{\lambda\mu}\psi. \tag{4.17}$$

We use also

$$G^\sigma{}_\lambda\equiv R^\sigma{}_\lambda-\tfrac{1}{2}\delta^\sigma{}_\lambda R\equiv\tfrac{1}{4}\eta^{\sigma\alpha\delta\beta}\eta_{\delta\mu\nu\lambda}R_{\alpha\beta}{}^{\mu\nu},$$

and we obtain

$$\nabla_\alpha E^{\sigma\alpha}=-4\nabla_\alpha\bar\psi(\gamma^\sigma\gamma^{\alpha\beta}+\gamma^{\alpha\beta}\gamma^\sigma)\nabla_\beta\psi+2G^\sigma{}_\lambda\bar\psi\gamma^\lambda\psi. \tag{4.18}$$

Einstein's equations, $G^\sigma{}_\lambda=T^\sigma{}_\lambda$, can be used to write the second term as

$$2T^\sigma{}_\lambda u^\lambda.$$

We know that $T^\sigma{}_\lambda n_\sigma u^\lambda$ is non-negative if n is time-like future-directed, u time-like past-directed, and the sources satisfy the dominant energy condition. We shall now study the contribution of the first term in the integral (4.17), when Σ is the space-like manifold $x^0=0$. It reduces then to (one has $\gamma^0\gamma^{0i}+\gamma^{0i}\gamma^0=0$)

$$\begin{aligned}2\nabla_i\bar\psi(\gamma^0\gamma^{ij}+\gamma^{ij}\gamma^0)\nabla_j\psi&=4\nabla_i\bar\psi\gamma^{ij}\nabla_j\psi\\&=4g^{ij}\nabla_i\bar\psi\nabla_j\psi+4\nabla_i\bar\psi\gamma^i\gamma^j\nabla_j\psi.\end{aligned} \tag{4.19}$$

We have proved the identity (valid if $G^0{}_\lambda$ is integrable on Σ)

$$U^\lambda P_\lambda=\int_S\omega=\int_\Sigma 2(g^{ij}\nabla_i\bar\psi\nabla_j\psi+G^0{}_\lambda\bar\psi\gamma^\lambda\psi+\nabla_i\bar\psi\gamma^i\gamma^j\nabla_j\psi)\mathrm{d}\Sigma. \tag{4.20}$$

We deduce from eq. (4.20) that $U^\lambda P_\lambda\geqslant 0$ if ψ satisfies the Witten equation

$$\mathscr{D}\psi=\gamma^j\nabla_j\psi=0. \tag{4.21}$$

Identity (4.20) gives, when ψ satisfies eq. (4.21) and $G_{\lambda\mu}=T_{\lambda\mu}$, a manifestly positive expression for $P_\lambda U^\lambda$, which shows that P^λ is time-like or null, and future-directed, since U^λ is arbitrary, time-like or null and

past-directed:

$$P_\lambda U^\lambda = 2\int_\Sigma (|\nabla\psi|^2 + T^\sigma{}_\lambda n_\sigma u^\lambda)\mathrm{d}\Sigma. \tag{4.22}$$

Note that the integrand in eq. (4.22) does not give a positive density of local energy for the gravitational field since ψ, solution of eq. (4.21), is a non-local functional of the initial data (g, k).

4.3. Solution of the Witten equation

Theorem. *Given asymptotically Minkowskian* H^2_{-1} (*cf. definition* 3, § 1) *initial data on* Σ, *such that* $G^0{}_\lambda$ *is integrable on* Σ *and satisfies the dominant energy condition* $G^0{}_\lambda u^\lambda \geqslant 0$ *for* u *non space-like, the Witten equation* (4.21) *admits one, and only one, solution* ψ *such that*

$$\psi = \psi_0 + \psi_1,$$

where ψ_0 *is an arbitrary given smooth spinor, constant in* $\Sigma \setminus K$, *and* $\psi_1 \in H^2_{-1}$.

Proof. We show that the operator $\gamma^j\nabla_j$ on the fields of spinors on Σ is an elliptic operator, injective from the space H^2_{-1} into the space H^1_0.

(a) The operator is elliptic since its principal symbol $\gamma^j\xi_j$ is an invertible matrix for each $\xi = (\xi_j) \neq 0$ (its square is, up to the positive factor $g^{ij}\xi_i\xi_j$, the unit 4×4 matrix).

(b) It is injective on H^2_{-1} because of the identity (4.20) which for an arbitrary spinor $\psi = \psi_0 + \psi_1$, $\psi_1 \in H^2_{-1}$, reads

$$\int_{\partial\Sigma} \omega = \int_\Sigma \mathrm{d}\omega = 4\int_\Sigma (g^{ij}\widetilde{\nabla_i\psi}\nabla_j\psi + \widetilde{\nabla_i\psi}\gamma^i\gamma^j\nabla_j\psi + \tfrac{1}{2}G^0{}_\lambda\bar{\psi}\gamma^\lambda\psi)\mathrm{d}\Sigma.$$

and implies if $\psi_0 = 0$ that $\int_{\partial\Sigma}\omega = 0$. Thus if $\psi \in H^2_{-1}$ and $\gamma^j\nabla_j\psi = 0$, by the positivity of the integrand, $\nabla_j\psi = 0$.

Lemma. If $\nabla_j\psi = 0$ *on* Σ *and* $\psi \in H^2_{-1}$ *then* $\psi = 0$ *on* Σ.

Proof of lemma. The norm of a spinor ψ is defined by

$$|\psi|^2 = \psi\bar{\psi} = -\bar{\psi}\gamma^0\psi; \tag{4.23}$$

it is a scalar if one remembers that in this equality γ^0 is the expression in the orthonormal frame with time axis the unit normal n of the

1-spinor − 1 cospinor $\gamma^\lambda n_\lambda$. Thus

$$\nabla_i|\psi|^2 = -\nabla_i\bar{\psi}\gamma^0\psi - \bar{\psi}\gamma^0\nabla_i\psi - \bar{\psi}\gamma^\lambda\nabla_i n_\lambda\psi. \tag{4.24}$$

If $\nabla_i\psi = 0$, then also $\nabla_i\bar{\psi} = 0$ and therefore

$$\nabla_i|\psi|^2 = -\bar{\psi}\gamma^\lambda\nabla_i n_\lambda\psi = -\bar{\psi}\gamma^j\omega_{ij}\psi = -\bar{\psi}\gamma^j k_{ij}\psi, \tag{4.25}$$

which implies the inequality, for the gradient of $|\psi|^2$, at each point of Σ,

$$|\nabla|\psi|^2| \leqslant c|k||\psi|^2, \tag{4.26}$$

where c is some fixed number and $|k|$ the norm of the extrinsic curvature k_{ij} of Σ. One shows that if $\psi \in H^2_{-1}$ the inequality (4.26) implies $\psi = 0$, by a method inspired by the one used by Christodoulou and O'Murchadha (1981) to prove the non-existence of Killing vector fields vanishing at infinity in asymptotically Euclidean manifolds.

From (a) and (b) it results (cf. Choquet-Bruhat and Christodoulou 1981) that $\mathcal{D} = \gamma^i\nabla_i$ is an isomorphism from H^2_{-1} onto H^1_0, that is $\mathcal{D}\psi = f$ has one, and only one, solution $\psi \in H^2_{-1}$ for every $f \in H^1_0$.

The theorem is proved by taking $f = -\mathcal{D}\psi_0$, $\psi = \psi_1$.

Hilbert space methods (cf. for instance an elementary exposition in Choquet-Bruhat et al. 1982, p. 507) can also be used (cf. Parker and Taubes 1982, Reula 1982) to prove the existence.

4.4. Zero energy implies flatness of space–time

The energy P^0 is zero if and only if the energy–momentum vector is zero, since it is non space-like. Witten proves that $P = 0$ implies that the Riemann curvature of ${}^{(4)}g$ is zero. His proof extends to the case of several ends: if $P_A = 0$ in one end Ω_A, then the space–time is flat. It proceeds as follows.

The identity (4.21) shows that $P_A = 0$ implies that the space–time is empty and every solution $\psi = \psi_0 + \psi_1$ of Witten's equation $\mathcal{D}\psi = 0$ is moreover such that

$$\nabla_i\psi = 0. \tag{4.26}$$

By the Ricci identity

$$R_{ij,\alpha\beta}\gamma^\alpha\gamma^\beta\psi = 0, \tag{4.27}$$

one deduces from (4.26):

Lemma. *If*

(1) *the spinors* $\psi_0{}^{(J)}$, $J = 1, \ldots, 4$, *are linearly independent in the end* Ω_A, *that is, there exists no set* a_J *of all nonzero numbers such that in* Ω_A

$$\sum_{J=1}^{4} a_J \psi_0{}^{(J)} \equiv 0.$$

(2) $\psi^{(J)} = \psi_0{}^{(J)} + \psi_1{}^{(J)}$ *with* $\psi_0{}^{(J)}$ *constant in* Ω_A, *zero in* Ω_B, $B \not\equiv A$, *and smooth,* $\psi_1{}^{(J)} \in H^2_{-1}$, *satisfy eq.* (4.26).

Then the $\psi^{(J)}$ *form a linearly independent set at each point* $x \in \Sigma$.

Proof. If $\sum_J a_J \psi^{(J)}(x) = 0$ at a point $x \in \Sigma$ one proves, using (4.26), that $\sum_J a_J \psi^{(J)} = 0$ on Σ, which is incompatible with hypothesis (1) and $\psi^{(A)} - \psi^{0(A)} \in H^2_{-1}$.

It is always possible to choose a complete linearly independent set of constant spinors $\psi_0{}^{(J)}$ in Ω_A, thus if $P_A = 0$ there is at each point $x \in \Sigma$ a complete linearly independent set of spinors ψ satisfying eq. (4.26), and thus eq. (4.27). Equation (4.27) therefore implies $R_{ij,\alpha\beta} = 0$.

Witten uses then the "many-fingered" character of time, that is the possibility to deform the hypersurface Σ in M (leaving it asymptotically fixed), to conclude that the full Riemann tensor $R_{\alpha\beta,\lambda\mu}$ vanishes on M.

Ashtekar and Horowitz (1982) have proved an interesting result which agrees with our physical intuition of asymptotically Minkowskian metrics as representing isolated systems. The ADM energy–momentum is in fact strictly time-like for non-flat space–time: the hypothesis that P is a null vector—which was not ruled out by the positive-energy theorem—implies that P is zero. Their proof uses the conformal treatment of space-like infinity. Their result can probably* be obtained directly, using the fact (Taub 1983) that the only space–times which admit covariantly constant spinors have a Weyl tensor algebraically special and null and, in addition, a Ricci tensor of the radiative type, that is,

$$R_{\alpha\beta} = -\tfrac{1}{2}\xi l_\alpha l_\beta.$$

5. Recent developments

5.1. *Positive-mass theorem for black holes*

The second paper of Schoen and Yau (1981b) proving the positivity of energy includes already the case where singularities are present, but hidden behind horizons. Using the Witten method, Gibbons et al. (1983a) give a

* It has now been proved (Taub 1984).

direct proof of this fact. They also prove that for a charged black hole one has the inequality

$$m^2 \geqslant Q^2 + P^2, \tag{5.1}$$

where Q and P are the electric and magnetic charges of the black hole, while, as before, $m^2 = -P^\mu P_\mu$, with P_μ the ADM four-momentum. We sketch the proof of the positivity of m.

Let Σ be a space-like submanifold with boundary of the space–time $(M, {}^{(4)}g)$, which is asymptotically Euclidean, but now Σ has moreover a boundary H, which is included in a large enough K. For an exact differential $d\omega$ we now have, where S denotes the "integral at infinity" previously defined,

$$\int_\Sigma d\omega = \int_{\partial\Sigma} \omega = \int_S \omega + \int_H \omega. \tag{5.2}$$

If ω is the Witten two-form (4.25), with $\psi = \psi_0 + \psi_1$, ψ_0 constant in $\Sigma \setminus K$, $\psi_1 \in H^1_{-1}$, and ψ satisfies the Witten equation $\not{D}\psi = 0$, we have as in § 4 ($\Sigma : x^0 = 0$)

$$4 \int_\Sigma (g^{ij} \tilde{\nabla}_i \bar{\psi} \nabla_j \psi + T^0_\lambda \bar{\psi} \gamma^\lambda \psi) d\psi = P_\lambda U^\lambda + \int_H \omega, \tag{5.3}$$

where P_λ is the ADM energy-momentum defined by integrals over S, two-spheres at infinity in the ends Ω_A, of an asymptotically Euclidean space. If ψ can be chosen such that $\int_H \omega = 0$, then the positivity of $P_\lambda U^\lambda$ is insured. It cannot be required that $\psi = 0$ on H, since the equation $\not{D}\psi = 0$ would then imply that the derivative of ψ transversal to H also vanishes, so that ψ is zero everywhere. Gibbons et al. (1983a) choose the boundary condition

$$\gamma^1 \gamma^0 \psi - \psi = 0 \quad \text{on} \quad H, \tag{5.4}$$

where e_1 denotes the unit vector orthogonal to H (in the tangent plane to Σ, e_0 is always the unit vector orthogonal to Σ). The condition (5.4) restricts the freedom of ψ on H by half, since the matrix $\gamma^1\gamma^0$ has the eigenvalues ± 1, with eigenspaces of dimension 2. Assuming that the Witten equation admits a solution ψ satisfying the previous asymptotic requirements and this boundary condition (which is likely), they show that the integral of ω on H vanishes, by rewriting it in the form (after

some algebraic manipulations using also $\mathcal{D}\psi = 0$)

$$-2\int_H \bar{\psi}((J+k-k_{11})\gamma^1\gamma^0 + 2\gamma^1\gamma^A D_A)\psi \mathrm{d}H, \tag{5.5}$$

where one has set ($A = 2, 3$)

$$D_A \equiv {}^{(2)}\nabla_A\psi + \tfrac{1}{2}k_{A1}\gamma^1\gamma^0\psi, \qquad {}^{(2)}\nabla_A \equiv \partial_A - \tfrac{1}{4}\omega_A{}^B{}_C\gamma_B\gamma^C,$$

and where J denotes the mean extrinsic curvature of H in Σ. The operator $L \equiv \gamma^1\gamma^A D_A$ anticommutes with $\gamma^1\gamma^0$, thus the last term of (5.5) vanishes if ψ satisfies (5.4), since then also

$$\bar{\psi}\gamma^1\gamma^0 = \bar{\psi}$$

and

$$\bar{\psi}L\psi = \bar{\psi}\gamma^1\gamma^0 L\psi = -\bar{\psi}L\gamma^1\gamma^0\psi \equiv -\bar{\psi}L\psi,$$

which implies $\bar{\psi}L\psi = 0$.

If H is an apparent horizon we have (cf. Hawking and Ellis 1973)

$$J + k - k_{11} = 0.$$

Thus (5.5) is zero, and the positivity of $P_\lambda U^\lambda$ is insured.

The inequality (5.1) is proved by replacing the covariant derivative of a spinor by (cf. Gibbons et al. 1983a)

$$\hat{\nabla}_\mu\psi = \nabla_\mu\psi + \tfrac{1}{4}\mathrm{i}F_{\alpha\beta}\gamma^\alpha\gamma^\beta\gamma_\mu\psi.$$

5.2. *The supergravity origin of Witten's expression*

It has been remarked very early in the development of supergravity (Deser and Teitelboim 1977) that the commutator of two supersymmetry transformations is a Lie derivative of the tetrad, and thus also of the metric, that is, if

$$\delta_\alpha e_a^\lambda = \bar{\alpha}\gamma_a\psi^\lambda, \qquad \delta_\alpha\psi_\lambda = D_\lambda\alpha, \tag{5.6}$$

we have, if α is an anticommuting spinor

$$(\delta_{\alpha_2}\delta_{\alpha_1} - \delta_{\alpha_1}\delta_{\alpha_2})e_a^\lambda = \nabla_a\xi^\lambda, \quad \text{with } \xi^\lambda = \bar{\alpha}_1\gamma^\lambda\alpha_2, \tag{5.7a}$$

$$(\delta_{\alpha_2}\delta_{\alpha_1} - \delta_{\alpha_1}\delta_{\alpha_2})g_{\lambda\mu} = \nabla_\lambda\xi_\mu + \nabla_\mu\xi_\lambda = (L_\xi g)_{\lambda\mu}. \tag{5.7b}$$

Thus, in some sense, the gravitational energy, "generator of time translations at infinity", should be the square of supercharge, generating super-

symmetry transformation. In fact, in quantum supergravity the Hamiltonian H may be written, at least formally

$$H = (1/\hbar)\mathrm{tr}\, Q^2, \tag{5.8}$$

where Q is the supercharge. It has also been suggested (Grisaru 1978) that it might be possible to obtain the positive-energy theorem in General Relativity by taking a classical limit, $\hbar \to 0$, of eq. (5.8). These statements have not been proved (cf. comments in Witten 1981), but since the appearance of Witten's proof, its link with classical supergravity theory has been clarified, in the Hamiltonian formalism with constraints (Horowitz and Strominger 1983, Deser 1983) and in the covariant formalism (Hull 1983). We give below a very simple derivation of Witten's expression, which, in accordance with eq. (5.7) (and in the spirit of § 1.1) is the quadratic form $\delta_\alpha Q^{(\alpha)}$, where Q_α is the conserved charge of the gravitino: we take here the spinor field α to be an arbitrary Dirac commuting (numerical valued) spinor—in fact, from supergravity we shall only use the definitions (5.6) of the supersymmetry transformation.

The Rarita–Schwinger equation for a spin-$\frac{3}{2}$ field ψ_λ, in a space–time $(M, {}^{(4)}g)$, is

$$B^\lambda({}^{(4)}g, \psi) \equiv \eta^{\lambda\mu\nu\rho}\zeta\gamma_\mu D_\nu\psi_\rho = 0. \tag{5.9}$$

Its linearization at ${}^{(4)}g = {}^{(4)}\hat{g}$, $\psi_\rho = 0$, is the linear operator acting on $\delta\psi_\rho = \varphi_\rho$,

$$B'^\lambda_\psi({}^{(4)}\hat{g}, 0)\cdot\varphi \equiv \hat{\eta}^{\lambda\mu\nu\rho}\hat{\gamma}_\mu\hat{\nabla}_\nu\varphi_\rho, \tag{5.10}$$

where $\hat{\nabla}_\nu$ is the ordinary covariant derivative. We have identically, for an arbitrary Dirac spinor φ_ρ (commuting or anticommuting is irrelevant, since eq. (5.10) is linear in φ_ρ),

$$\hat{\nabla}_\lambda(\hat{\eta}^{\lambda\mu\nu\rho}\zeta\hat{\gamma}_\mu\hat{\nabla}_\nu\varphi_\rho) \equiv 2\hat{G}^{\lambda\rho}\gamma_\lambda\varphi_\rho, \qquad \hat{G}^{\lambda\rho} = \hat{R}^{\lambda\rho} - \tfrac{1}{2}\hat{g}^{\lambda\rho}R, \tag{5.11}$$

(which can also be deduced from the fundamental identity of supergravity $D_\lambda B^\lambda \equiv 2(G^{\lambda\rho} - T^{\lambda\rho})\gamma_\lambda\psi_\rho$, but only for Majorana anticommuting spinors). We suppose that the background ${}^{(4)}\hat{g}$ is a solution of empty-space Einstein equations

$$\hat{G}_{\lambda\rho} = 0.$$

We suppose also that ${}^{(4)}\hat{g}$ admits a Killing, that is, covariantly constant, spinor α, thus

$$\hat{\nabla}_\lambda\bar{\alpha} = 0 \quad \text{and also} \quad \hat{\nabla}_\lambda\alpha = 0.$$

We deduce then from eq. (5.11) the conservation law

$$\hat{\nabla}_\lambda q^\lambda = 0,$$

where q^λ is the vector field

$$q^\lambda = \bar{\alpha}\hat{\eta}^{\lambda\mu\nu\rho}\hat{\gamma}_\mu\hat{\nabla}_\nu\varphi_\rho.$$

In accordance with general principles (cf. DeWitt 1984) and § 1.1, the charge $Q(\alpha)$ associated with the Killing spinor α and a slice Σ of M, for a spin-$\frac{3}{2}$ field φ_ρ in the space-time $(M, {}^{(4)}g)$, relative to the configuration $(M, {}^{(4)}\hat{g}, 0)$, is

$$Q(\alpha) = \int_\Sigma \bar{\alpha}\hat{\eta}^{\lambda\mu\nu\rho}\xi\hat{\gamma}_\mu\hat{\nabla}_\nu\varphi_\rho \, d\Sigma_\lambda, \qquad d\Sigma_\lambda = \frac{1}{3!}\hat{\eta}_{\lambda\alpha\beta\gamma}dx^\lambda \wedge dx^\beta \wedge dx^\gamma. \tag{5.12}$$

This charge is conserved if appropriate asymptotic behavior is required on φ_ρ. It can also be written as an integral on $\partial\Sigma$, since the integrand is an exact differential,

$$Q(\alpha) = \int_{\partial\Sigma} \bar{\alpha}\xi(\hat{\gamma}_\mu\varphi_\rho - \hat{\gamma}_\rho\varphi_\mu)dx^\mu \wedge dx^\rho. \tag{5.13}$$

If ${}^{(4)}g = {}^{(4)}\hat{g}$ on $\partial\Sigma$ then $Q(\alpha)$ can equivalently be written

$$Q(\alpha) = \int_{\partial\Sigma} \bar{\alpha}\xi(\gamma_\mu\varphi_\rho - \gamma_\rho\varphi_\mu)dx^\mu \wedge dx^\rho; \tag{5.14}$$

by a supersymmetry transformation with parameter α, at $\psi = 0$ we obtain

$$\delta_\alpha Q(\alpha) = \int_{\partial\Sigma} \bar{\alpha}\xi(\nabla_\mu\nabla_\rho\alpha - \gamma_\rho\nabla_\mu\alpha)dx^\mu \wedge dx^\rho,$$

which is precisely Witten's two-form.

5.3. Asymptotically anti-De Sitter space-times (Abbott and Deser 1982, Gibbons et al. 1983a, b)

An anti-De Sitter space-time is believed to be the ground state—because maximally supersymmetric—of gauged supergravity.

An anti-De Sitter (aDS) space-time is the universal covering space of the quadric C in $\mathbb{R}^5$, invariant by O(3, 2), given by

$$-(X^0)^2+(X^1)^2+(X^2)^2+(X^3)^2-(X^4)^2 = 3\Lambda^{-1},$$

where $\Lambda < 0$, with metric induced from flat space with signature $(-+++-)$. It has topology $\mathbb{R}^4$ and its metric ${}^{(4)}\hat{g}$ satisfies Einstein equations with cosmological constant

$$\hat{R}_{\lambda\mu} = \Lambda \hat{g}_{\lambda\mu}.$$

One may choose coordinates $t \in \mathbb{R}$, r, θ, φ, standard polar coordinates in $\mathbb{R}^3$, in which the metric takes the static form:

$$-\cosh^2(\sqrt{\tfrac{1}{3}|\Lambda|}r)\mathrm{d}t^2+\mathrm{d}R^2+\sinh^2(\sqrt{\tfrac{1}{3}|\Lambda|}r)(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2).$$

The time-like Killing vector $\xi^\mu = (-1, 0)$ can be used to define an energy-momentum for asymptotically aDS space-times, relative to aDS. This energy has been studied by the methods of §§ 2 and 4, and though the proofs published are not complete in all mathematical details, the results are in agreement with positivity.

References

Abbott, L., and S. Deser, 1982, Nucl. Phys. **B195**, 76–96.

Arnowitt, R., S. Deser and C. Misner, 1962, in: Gravitation, an Introduction to Current Research, ed. L. Witten (Wiley, New York) p. 227.

Ashtekar, A., and G.W. Horowitz, 1982, Phys. Lett. **89A**, 181–183.

Atiyah, M., and R. Bott, 1983, IHES report.

Avez, A., 1965, Ann. Inst. Fourier **13**, 105–190.

Bartnik, R., 1984, in: Proc. Symp. on Asymptotic properties of spacetime, ed. F. Flaherty (Springer, Berlin), to appear.

Brill, D., and S. Deser, 1968, Ann. Phys. **50**, 548.

Brill, D., and Soo Yang Pong, 1980, The Positive Mass Conjecture, in: General Relativity and Gravitation, ed. A. Held (Plenum, New York).

Chaljub-Simon, A., and Y. Choquet-Bruhat, 1979, Ann. Univ. Toulouse, no. **1**, 25–29.

Choquet-Bruhat, Y., and G. Choquet, 1978, C.R. Acad. Sci. Paris **287**, 1047.

Choquet-Bruhat, Y., and D. Christodoulou, 1981, Acta Math. **146**, 129–150.

Choquet-Bruhat, Y., and J. Marsden, 1976, Commun. Math. Phys. **51**, no. 3, 283.

Choquet-Bruhat, Y., and T. Ruggeri, 1983, Commun. Math. Phys. **89**, 269–275.

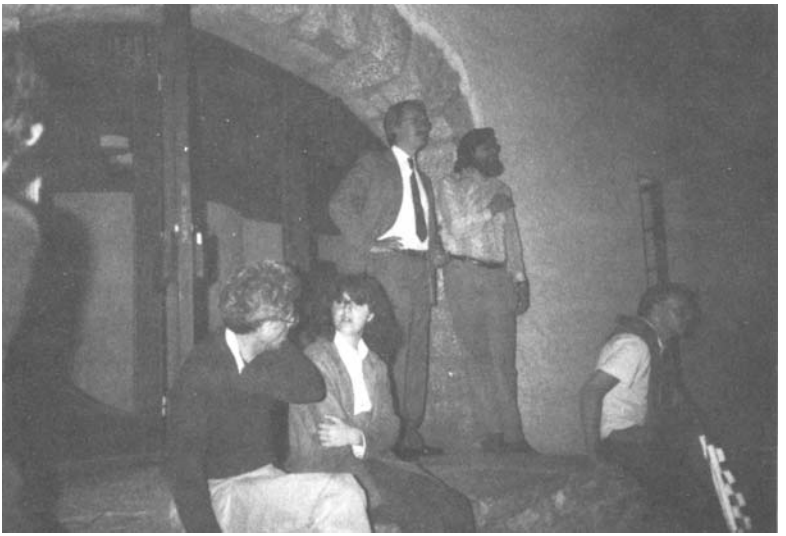
Choquet-Bruhat, Y., and J.W. York, 1980, The Cauchy problem, in: General Relativity and Gravitation, ed. A. Held (Plenum, New York).

Choquet-Bruhat, Y., A. Fischer and J. Marsden, 1977, C.R. Acad. Sci. Paris **284**, 975.

Choquet-Bruhat, Y., A. Fischer and J. Marsden, 1979, Maximal Surfaces and Positivity of Mass, in: Isolated Gravitating Systems in General Relativity, ed. J. Ehlers (North-Holland, Amsterdam).

Choquet-Bruhat, Y., and C. DeWitt-Morette, with M. Dillard-Bleick, 1982, Analysis Manifolds and Physics, 2nd Rev. Ed. (North-Holland, Amsterdam).

Christodoulou, D., and N. O'Murchadha, 1981, Commun. Math. Phys. **80**, 277.
Deser, S., 1983, Phys. Rev. **D27**, 2805.
Deser, S., and L.D. Faddeev, 1968, Phys. Lett. **26A**, 538.
Deser, S., and C. Teitelboim, 1977, Phys. Rev. Lett. **39**, 249.
DeWitt, B.S., 1967, The Dynamical Theory of Groups and Fields (Gordon and Breach, New York).
DeWitt, B.S., 1984, The Spacetime Approach to Quantum Field Theory, this volume.
Fischer, A., and J. Marsden, 1979, Topics in the Dynamics of General Relativity, in: Isolated Gravitating Systems in General Relativity, ed. J. Ehlers (North-Holland, Amsterdam).
Fischer, A., and J. Marsden, 1979, in: General Relativity, eds. S.W. Hawking and W. Israel (Cambridge Univ. Press).
Geroch, R., 1973, Ann. N.Y. Acad. Sci. **224**, 108.
Gibbons, G.W., S.W. Hawking, G.W. Horowitz and M.J. Perry, 1983a, Commun. Math. Phys. **88**, 295.
Gibbons, G.W., N. Hull and N. Warner, 1983, to be published.
Grisaru, M.T., 1978, Phys. Lett. **73B**, 207.
Hawking, S.W., and G.F.R. Ellis, 1973, The Large Scale Structure of Space-Time (Cambridge Univ. Press).
Horowitz, G.W., and Strominger, 1983, Phys. Rev. **D27**, 2793.
Hull, N., 1983, DAMTP preprint (Univ. of Cambridge).
Kazdan, J., 1982, Séminaire Bourbaki, Paris.
Kazdan, J., and F. Warner, 1975, Proc. Symp. Pure Math. (Amer. Math. Soc.) **27**, 219–226.
Lichnerowicz, A., 1944, J. Math. Pures Appl. **23**, 37–63.
Lichnerowicz, A., 1961, Publ. IHES, no. **11**.
Lichnerowicz, A., 1964, Ann. Inst. H. Poincaré t. **1**, no. 3, 233–290.
Nester, J., 1981, Phys. Lett. **83A**, 241–242.
O'Murchadha, N., and J. York, 1974, Phys. Rev. **D10**, 2–345.
Parker, T., and T. Taubes, 1982, Commun. Math. Phys. **84**, 223–238.
Pong, Soo Yang, 1976, J. Math. Phys. **17**, 141.
Pong, Soo Yang, 1978, J. Math. Phys. **19**, 1155.
Regge, T., and C. Teitelboim, 1974, Ann. Phys. **88**, 286–318.
Reula, O., 1982, J. Math. Phys. **23**, 810–813.
Schoen, R., and S. Yau, 1979, Commun. Math. Phys. **65**, 45–76.
Schoen, R., and S. Yau, 1981a, Commun. Math. Phys. **79**, 231–260.
Schoen, R., and S. Yau, 1981b, Commun. Math. Phys. **79**, 47–51.
Taub, A., 1983, private communication.
Taub, A., 1984, Ann. Inst. H. Poincaré, to be published.
Witten, E., 1981, Commun. Math. Phys. **80**, 381.
Witten, E., 1981, Nucl. Phys. **B195**, 481.
York, J.W., 1979, Kinematics and Dynamics of General Relativity, in: Sources of Gravitational Radiation, ed. L. Smarr (Cambridge Univ. Press).
York, J.W., 1980, Energy and Momentum of the Gravitational Field, in: Essays in General Relativity, ed. F. Tipler (Academic, New York).



REFERENCES

Abraham, A., Anderson, A., Choquet-Bruhat, Y., and York, J. (1995) *Phys. Rev. Lett.*, **75**, 3337–81.

Anderson, A. and York, J. (1999) *Phys. Rev. Lett.*, **82**, 4384–7.

Anderson, A., Choquet-Bruhat, Y. and York, J. (1997) Einstein–Bianchi hyperbolic system for General Relativity, *Topol. Meth. Non-Linear Anal.*, **10**(2), 353–73.

Anderson, M. (2001) *Commun. Math. Phys.*, **222**(3), 533–67.

Anderson, M. and Chruściel, P. (2005) *Commun. Math. Phys.*, **260**, 557–77.

Andersson, L. and Chruściel, P. (1994), *Commun. Math. Phys.*, **161**(3), 533–68.

Andersson, L. and Moncrief, V. (2001) *Ann. H. Poincaré*, **4**(1), 1–34.

Andersson, L. and Moncrief, V. (2004) in *The Einstein Equations and the Large Scale Behavior of Spacetime* (eds. P. Chruściel and H. Friedrich) Birkhaüser, pp. 299–330.

Andersson, L. and Rendall, A. (2001) *Commun. Math. Phys.*, **218**, 479–511.

Andersson, L., Chruściel, P., and Friedrich, H. (1992) *Commun. Math. Phys.*, **149**, 587–612.

Andersson, L., Moncrief, V., and Tromba, A. (1997) *J. Geom. Phys.*, **23**(3,4), 1991–2005.

Andréasson H. 2005 arXiv gr-qc/0502091 v1.

Anile, M. (1987) *Relativistic Hydrodynamics and Magnetohydrodynamics*, Cambridge University Press.

Anile, M. and Greco, A. (1979) *Nuovo Cimento B*, **55**(2), 307–24.

Arnowitt, R., Deser, S., and Misner, C. W. (1962) In *Gravitation, an Introduction to Current Research* (ed. L. Witten), Wiley.

Ashtekar, A. and Dryer, O. (2003) gr-qc 0206024 V3.

Aubin, T. (1970) *J. Diff. Geom.*, **4**, 383–424.

Aubin, T. (1982) *Non-Linear Analysis and Monge–Ampère Equations*, Springer, London.

Aubin, T. (1976) *Math. Pures Appl.*, **55**, 269–96.

Bancel, D. (1970) *Ann. Inst. Poincaré* (Boltzmann equation)

Bancel, D. and Lacaze, J. (1976) *C. R. Acad. Sci. Paris*, **283**, 405–7.

Bancel, D. and Choquet-Bruhat, Y. (1973) *Commun. Math. Phys.*, **33**, 1–14.

Baouendi, S. and Goulaouic, C. (1977) *Commun. PDE*, **2**, 1151–62.

Bär, C., Ginoux, N., and Pfäffle (2007) Wave equations on Lorentzian manifolds and quantization. *Eur. Math. Soc.*, ETH Zurich.

Bartnik, R. (1984) *Commun. Math. Phys.*, **94**, 155–75.

Bartnik, R. (1986) *Commun. Pures Appl. Math.*, **39**, 661–93.

Bartnik, R. (1988) *Commun. Math. Phys.*, **117**, 615–24.

Beig, R. (1978) *Phys. Rev. Lett. A*, **69**(3), 153–5.

Beig, R. and Simon, W. (1981) *Proc. R. Soc. Lond. Ser. A*, **376**, 333–41.

Beig, R. and Chruściel, P. (2006) gr-qc 0612012v2.

Beig, R., Chruściel, P., and Schoen, R. (2005) *Ann. H. Poincaré*, **6**, 428–46.

Bel, L. (1958) *C. R. Acad. Sci. Paris.*

Belinskii, V. A. and Khalatnikov, I. M. (1973) *Sov. Phys. JETP*, **36**, 591–7.
Belinski, V. A., Khalatnikov, I. M., and Lifshitz, E. M. (1970) *Adv. Phys.*, **19** 525–73.
Belinskii, V. A., Lifshitz, E. M., and Khalatnikov, I. M. (1971) *Sov. Phys. Usp.*, **13**, 745.
Belinskii, V. A., Lifshitz, E. M., and Khalatnikov, I. M. (1972) *Sov. Phys. JETP*, **35**, 838.
Berger, B. and Moncrief, V. (1993) *Phys. Rev. D*, **48**(10), 4676–87.
Berger, B. and Moncrief, V. (1998) *Phys. Rev. D*, **58**, 064023.
Berger, B., Chruściel, P., Isenberg, J., and Moncrief, V. (1997) *Ann. Phys.*, **260**, 117–48.
Berger, B. Isenberg, J., and Weaver, M. (2004) *Phys. Rev. D*, **64**, 84–6.
Boillat, G. (1996) in *Recent Mathematical Methods in Non Linear Wave Propagation*, Springer Lecture Notes in Maths, 1640.
Boillat, G. and Ruggeri, T. (1997) *Arch. Rat. Mech. Anal.* **137** , 305–20.
Boillat, G. and Ruggeri, T. (1999) *Cont. Mech. Thermodyn.*, **11**, 107–11.
Boillat, G. and Ruggeri, T. (1999) *J. Math. Phys.*, **40**(12), 6399–404.
Bondi, H. (1947) *Mon. Nat. R. Astron. Soc.*, **107**, 410–25.
Bott, R. (1968) *Batelle Rencontres*, (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.
Bray, H. and Chruściel, P. (2004) The Penrose inequality. In *The Einstein Equations and the Large Scale Behaviour of Spacetime* (eds. P. Chruściel and H. Friedrich), Birkhäuser.
Brill, D. and Cantor, M. (1981) *Compos. Math.*, **43**, 317–25.
Brini, F. and Ruggeri, T. (1999) *Cont. Mech. Thermodyn.*, **11**, 331–8.
Cagnac, F. and Choquet-Bruhat, Y. (1984), *J. Math. Pures Appl.*, **63**, 377–90.
Carter, B. (1973) In *Black Holes* (eds. C. de Witt and B. de Witt), Gordon & Breach.
Cattaneo, C. (1959) *Ann. Math. Pura Appl.* **XLVIII**(IV), 361–86.
Chaljub-Simon, A. and Choquet-Bruhat, Y. (1979) *Ann. Univ. Toulouse*, **1**, 25–39.
Chandrasekhar, S. (1983) *The Mathematical Theory of Black Holes*, Clarendon Press, Oxford.
Chern, S. (1944) *Ann. Math.*, **45**, 747–52.
Choptuik, M. W. (1993) *Phys. Rev. Lett.*, **70**, 9.
Choquet-Bruhat, Y. (1958) *Bull. Soc. Math.*, **86**, 155–75.
Choquet-Bruhat, Y. (1960) *Astron. Acta*, **VI**(6), 354–65.
(Choquet-)Bruhat, Y. (1962) "The Cauchy problem" in *Gravitation: An Introduction to Current Research* (ed. L. Witten), pp. 130–68, Wiley.
Choquet-Bruhat, Y. (1962) *Annali di Matemetica*, serie 4, **64**, 191–228.
Choquet-Bruhat, Y. (1964) *C. R. Acad. Sci. Paris*, **258**, 3809–12
Choquet-Bruhat, Y. (1965) *C. R. Acad. Sci. Paris*, **261**, 354–6.
Choquet-Bruhat, Y. (1966) *Commun. Math. Phys.*, **3**, 334–57.
Choquet-Bruhat, Y. (1966) *J. Math. Pures Appl.*, **45**, 371–86.
Choquet-Bruhat, Y. (1967) in *Batelle Rencontres* (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.
Choquet-Bruhat, Y. (1968) *Géométrie Différentielle et Systèmes Extérieurs*, Dunod.
Choquet-Bruhat, Y. (1969) *J. Math. Pures Appl.*, **48**, 117–58.
Choquet-Bruhat, Y. (1969) *Commun. Math. Phys.*, **12**, 16–35.
Choquet-Bruhat, Y. (1971) *Commun. Math. Phys.*, **21**, 211–18.

Choquet-Bruhat, Y. (1972) *C. R. Acad. Sci. Paris*, **274**, 682–4.
Choquet-Bruhat, Y. (1972) *J. Gen. Rel. Grav.*, **2**(4), 359–62.
Choquet-Bruhat, Y. (1973) *C. R. Acad. Sci. Paris*, **276**A, 281–4.
Choquet-Bruhat, Y. (1976) *Ann. Scuola Norm. Pisa*, Serie IV No. 3, 361–78.
Choquet-Bruhat, Y. (1983) in *Relativity, Cosmology, Topological Mass and Supergravity* (ed. C. Aragone), pp. 108–35, World Scientific.
Choquet-Bruhat, Y. (1986) *C. R. Acad. Sci. Paris* serie I, **303**(4), 109–11.
Choquet-Bruhat, Y. (1987) *Ann. Inst. Poincaré*, **46**(1), 97–111.
Choquet-Bruhat, Y. (1989) *Class. Quant. Grav.*, **6**, 1781–6.
Choquet-Bruhat, Y. (1991) Marcel Grossmann meeting.
Choquet-Bruhat, Y. (1992) *J. Math. Phys.*, **33**(5), 1782–5.
Choquet-Bruhat, Y. (1992) *C. R. Acad. SciParis*, serie 1 **318**, 775–82.
Choquet-Bruhat, Y. (2000) *Ann. Phys. (Leipzig)*, **9**(3–5), 258–66.
Choquet-Bruhat, Y. (2004) *Class. Quant. Grav.*, **21**(3), 127–51.
Choquet-Bruhat, Y. (2004) in *Einstein Equations, Large Scale Behavior of Gravitational Fields* (eds. H. Friedrich and P. Chruściel), Birkhaüser.
Choquet-Bruhat, Y. (2005) Proc. Greek Relativity Meeting, Lesbos, *J. Phys. Conf. Ser.*, **8**, 1–12, gr-qc 0501052.
Choquet-Bruhat, Y. and Christodoulou, D. (1981) *Ann. Sci. E*; N.S. serie 4, **14**, 481–500.
Choquet-Bruhat, Y. and Christodoulou, D. (1981) *Acta Mathematica*, **146**, 146–50.
Choquet-Bruhat, Y. and Cotsakis, S. (2002) *J. Geom. Phys.*, **43**(4), 345–50.
Choquet-Bruhat, Y. and Deser, S. (1972) *C. R. Acad. Sci. Paris*, **274**, 682–4.
Choquet-Bruhat, Y. and DeWitt-Morette, C. (1982) *Analysis Manifolds and Physics I and II* (revised edition 2000).
Choquet-Bruhat, Y. and Friedrich, H. (2006) Class. Quant. Grav. 23, 5941–5949.
Choquet-Bruhat, Y. and Geroch, R. (1969) *Commun. Math. Phys.*, **14**, 329–35.
Choquet-Bruhat, Y. and Greco, A. (1983) *J. Math. Phys.*, **24**(2), 377.
Choquet-Bruhat, Y. and Gu Chao Hao (1989) *C. R. Acad. Sci. Paris*, serie 1, **308**, 167–70.
Choquet-Bruhat, Y. and Isenberg, J. (2006) *J. Geom. Phys.*, **6**, 1199–214.
Choquet-Bruhat, Y. and Leray, J. (1972) *C. R. Acad. Sci. Paris*, **274**, 81–5.
Choquet-Bruhat, Y. and Moncrief, V. (1996) Existence theorem for solutions of Einstein equations with 1 parameter spacelike isometry group. *Proc. Symp. Pure Math.*, **59** (eds. H. Brezis and I. E. Segal), pp. 67–80.
Choquet-Bruhat, Y. and Moncrief, V. (2001) Future global in time Einsteinian spacetimes with $U(1)$ isometry group. *Ann. Henri. Poincaré*, **2**, 1007–64.
Choquet-Bruhat, Y. and Noutchegueme, N. (1986) *C. R. Acad. Sci. Paris*, **303**, serie I 6, 259–63.
Choquet-Bruhat, Y. and Novello, M. (1987) *C. R. Acad. Sci. Paris*, **305**, 155–60.
Choquet-Bruhat, Y. and Ruggeri, T. (1983) *Commun. Math. Phys.*, **89**, 269–75.
Choquet-Bruhat, Y. and York, J. W. (1980) The Cauchy problem, in *General Relativity and Gravitation* (ed. A. Held), Plenum.
Choquet-Bruhat, Y. and York, J. W. (1996) in *Gravitation, Electromagnetism and Geometric Structures* (ed. Ferrarese), Pitagora.
Choquet-Bruhat, Y. and York, J. W. (1998) in *Gravitation, Electromagnetism and Geometric Structures* (ed. G. Ferrarese), Pitagora.

Choquet-Bruhat, Y., Chruściel, P., and Loiselet, J. (2006) *Class. Quant. Grav.*, **23**(24), 7383–94.
Choquet-Bruhat, Y., Isenberg, J., and Moncrief, V. (2005) *Il Nuovo Cimento B*, **119**(7–9).
Choquet-Bruhat, Y., Isenberg, J., and Pollack, D. (2006) *Chinese Ann. Math.*, **27**, serie B, No. 1, 31–52.
Choquet-Bruhat, Y., Isenberg, J., and Pollack, D. (2007) *Class. Quant. Grav.*, **24**(1), 800–29.
Choquet-Bruhat, Y., Isenberg, J., and York, J. (2000) *Phys. Rev. D*, **3**(8), 83–105.
Choquet-Bruhat, Y., Paneitz, S., and Segal, I. E. (1983) *J. Funct. Anal.*, **53**(2) 112–50.
Choquet, G. and Choquet-Bruhat, Y. (1978) *C. R. Acad. Sci. Paris*, **287A**, 1047–9.
Christodoulou, D. (1984) *Commun. Math. Phys.*, **93**, 171–95).
Christodoulou, D. (1986) *Commun. Math. Phys.*, **105**, 337–61.
Christodoulou, D. (1986) *Commun. Math. Phys.*, **106**, 587–621.
Christodoulou, D. (1986) *Commun. Pure Appl. Math.*, **39**, 267–82.
Christodoulou, D. (1987) *Commun. Math. Phys.*, **109**, 591–611.
Christodoulou, D. (1987) *Commun. Math. Phys.*, **109**, 613–47.
Christodoulou, D. (1981) *C. R. Acad. Sci. Paris*, **293**, 79–42.
Christodoulou, D. (1999) *Ann. Math.*, **149**, 143–217.
Christodoulou, D. (1999) *Class. Quant. Grav.*, **13**, A23–A35.
Christodoulou, D. (2000) *Marcel Grossmann Meeting on General Relativity*, MGIXMM, Gurzadyan (eds. Jantzen and Ruffini), pp. 44–54, World Scientific.
Christodoulou, D. (2007) *The Formation of Shocks in Relativistic Fluids*, EMS, Zurich.
Christodoulou, D. (2008) gr-qc 0805-3880.
Christodoulou, D. and Klainerman, S. (1993) *The Global Non-Linear Stability of Minkowski Space*, Princeton Mathematical series 41.
Christodoulou, D. and O'Murchada, N. (1981) *Commun. Math. Phys.*, **80**, 271–300.
Chruściel, P. (1986) *Class. Quant. Grav.*, L115–21.
Chruściel, P. (1992) *Contemp. Math.*, **132**, Amer. Math. Soc., Providence, RI.
Chruściel, P. (1999) *Class. Quant. Grav.*, **16**, 689–704, arXiv:gr-qc/98100.
Chruściel , P. (2000) "Introduction to Lorentzian geometry and Einstein equations in the large", Lectures at the Levoca Summer School.
Chruściel, P. (2006) gr-qc 0606101.
Chruściel, P. and Delay, E. (2002) *Class. Quant. Grav.*, **19**, L71–L79.
Chruściel, P. and Delay, E. (2002) gr-qc 0208092v1.
Chruściel, P. and Delay, E. (2003) Mémoires de la Société Mathématique de France, **94**, 1–103.
Chruściel, P. and Shatah, J. (1997) *Asian J. Math.*, **1**, 530–48.
Chruściel, P. and Tod, K. P. (2005) arXiv:gr-qc/0512043.
Chruściel, P. and Wald, R. (1994) *Class. Quant. Grav.*, **11** L147–52.
Chruściel, P., Beig, R., and Schoen, R. *Adv. Theor. Math. Phys.*, gr-qc/0403042.
Chruściel, P., Delay, E., Galloway, G., and Howard, J. (2001) *Ann. Inst. Henri Poincaré*, **2**(1), 109–78.
Chruściel, P., Isenberg, J., and Moncrief, V. (1990) *Class. Quant. Grav.*, **7**, 1671–80.
Chruściel, P., Isenberg, I., and Pollack, D. (2005) *Commun. Math. Phys.*, **256**, 29–47.
Claudel, C. (2000) gr-qc 0005031v2.

Corvino, J. (2000) *Commun. Math. Phys.*, **214**, 137–89.
Corvino, J. and Schoen, R. (2003) gr-qc 0301071.
Corvino, J. and Schoen, R. (2006) *J. Diff. Geom.*, **73**, 85–217.
Dafermos, C. (1994) Lecture Notes in Mathematics 1640 (ed. T. Ruggeri), Springer.
Dafermos, M. (2008) gr-gc 0307–013.
Dafermos, M. and Rendall, A. (2007) arXiv gr-qc 0610075v1.
Damour, T. "General Relativity Today", gr.qc 0704.0754v1.
Damour, T. (1986) *Fourth Marcel Grossmann Meeting on General Relativity*, (ed. Ruffini), pp. 365–92, Elsevier.
Damour, T. (2005) *Si Einstein m'était conté*, Cherche Midi, translated by E. Novak (2006) as *Once Upon Einstein*, A. K. Peters Ltd, Wellesley MA.
Damour, T. and de Buyl, S. *Describing General Cosmological Singularities in Iwasawa Variables*, to be published.
Damour, T. and Henneaux, M. (2000) *Phys. Lett. B*, **488**, 108.
Damour, T. and Henneaux, M. (2000) *Phys. Rev. Lett.*, **85**, 920.
Damour, T. and Taylor, J. H. (1991) *Astrophys. J.*, **366**, 501.
Damour, T., Henneaux, M., and Nicolai, H. (2003) *Class. Quant. Grav.*, **20**, R145–R200.
Damour, T., Henneaux, M., Rendall, A., and Weaver, M. (2002) *Ann. H. Poincaré*, **3**, 1049.
Darmois, G. (1923) *Les équations de la gravitation Einsteinienne*, Gauthier-Villars.
Demaret, J., Hanquin, J. L., Henneaux, M., Spindel, P., and Taormina, A. (1986) *Phys. Lett. B*, **175**, 129.
Demaret, J., Henneaux, M., and Spindel, P. (1985) *Phys. Lett. B*, **164**, 27.
DeWitt, C. and DeWitt, B. (eds.) (1983) *Relativity Groups and Topology*, Gordon and Breach.
Dionne, P. (1962) *J. d'Anal. Math. Jérusalem*, **1**, 1–90.
Dirac, P. A. M. (1959) *Phys. Rev.*, **114**, 924–30.
Douglis, A. and Nirenberg, L. (1955) *Commun. Pure Appl. Math.*, **8**, 503–38.
Eardley, D. and Moncrief, V. (1982) *Commun. Math. Phys.*, **83**, 171–92 and 193–203.
Eckart, C. (1940) *Phys. Rev.*, **58**, 919–24.
Eells, J. and Lemaire, L. (1974) *Bull. Math. Soc.*, **10**, 1–68.
Emparan, R. and Reall, H. S. (2002) *Phys. Rev. Lett.*, **88**, 101, arXiv:hep-th/0110260
Emparan, R. and Reall, H. S. (2004) *J. High Energy Phys.*, **3**, 64.
Emparan, R. and Reall, H. S. (2006) *Class. Quant. Grav.*, **23**, R169–R197, arXiv:hep-th/0608012.
Ferrarese, G. (1963) *Rendic. Matem.*, **22**, 147–68.
Ferrarese, G. (2004) *Riferimenti Generalizzati in Relatività e Applicazioni* Pitagora editrice, Bologna.
Ferrarese, G. and Bini, D. (2007) *Relativistic Mechanics of Continuous Media*, Springer.
Fisher, A. and Marsden, J. (1975) *Proc. Symp. Pure Math. AMS*, **27**, part 2, 219–63.
Fisher, A. and Moncrief, V. (1994) in *Physics on Manifolds* (eds. M. Flato, R. Kerner and A. Lichnerowicz), Kluwer, pp. 111–51.
Fisher, A. and Moncrief, V. (1994) in *Global Structures and Evolution in General Relativity* (eds. S. Cotsakis and G. Gibbons), pp. 111–73. Springer-Verlag, London.

Fisher, A. and Moncrief, V. (1998) in *Mathematical and Quantum aspects of General Relativity and Cosmology* (eds. S. Cotsakis and G. Gibbons), Lecture Notes in Physics 537, Springer, London, pp. 70–102.
Fisher, A. and Tromba, A. (1984) *Math. Ann.*, **167**, 118–42.
Fourès (Choquet)-Bruhat, Y. (1948) *C. R. Acad. Sci. Paris*, **226**, 48–51.
Fourès (Choquet)-Bruhat, Y. (1952) *Acta Mathematica*, **88**, 141–225.
Fourès (Choquet)-Bruhat, Y. (1956) *J. Rat. Mech. Anal.*, **5**, 951–66.
Fourès (Choquet)-Bruhat, Y. (1958) *Bull. Soc. Math. France*, **86**, 155–75.
Friedman, A. (1976) *Partial Differential Equations*, Part 1 Section 10, Robert Krieger Publishing Company.
Friedrich, H. (1981) *Proc. Roy. Soc. Lond.*, **A378**, 401–21.
Friedrich, H. (1985) *Commun. Math. Phys.*, **100**, 525–43.
Friedrich, H. (1986) *Commun. Math. Phys.*, **107**, 587–609.
Friedrich, H. (1996) *Class. Quant. Grav.*, **13**, 1451–9.
Friedrich, H. (1998) On the evolution equations for gravitating ideal fluid bodies in General Relativity. *Phys. Rev. D*, **57**, 2317–22.
Friedrich, H. (2004) in *The Einstein Equations and the Large Scale Behavior of Gravitational Fields* (eds. P. Chruściel and H. Friedrich), Birkhäuser.
Friedrichs, O. (1954) *Commun. Pure Appl. Math.*, **7**, 345–92.
Fritelli, S. and Reula, O. A. (1994) *Commun. Math. Phys.*, **166**, 221–35.
Galloway, G. (2002) gr-qc 0208092v1.
Garding, L., Kotake, T., and Leray, J. (1966) *Bull. Soc. Math. Fr.*, **94**, 25–48.
Garfinkle, D. (2007) *Class. Quant. Grav.*, **12**, S295–S306.
Geroch, R. (1970) *J. Math. Phys.*, **11** 437–49.
Gowdy, R. (1971) *Phys. Rev. Lett.*, **27**, 826–9.
Gowdy, R. (1974) *Ann. Phys.*, **83**, 203–41.
Gromov, M. and Lawson, H. B. (1980) *Ann. Math.*, **111**, 423–34.
Grubišić, B. and Moncrief, V. (1993) *Phys. Rev. D*, **47**(6), 2371–82.
Gu Chao Hao (1973) *J. Fudan University*, **1**, 73–8.
Gu Chao Hao (1980) *Commun. Pures Appl. Math.*, **33**, 727–37.
Gundlach, C., Calabrese, G., Hinder, I. and Martin-Garcial, J. (2005) *Class. Quant. Grav.*, **22**, 3767–774.
Hanouzet, B. and Natalini, R. (2003) *Arch. Rat. Mech. Anal.*, **169**, 89.
Hartle, B. and Hawking, S. W. (1972) *Commun. Math. Phys.*, **26**, 87–101.
Hawking, S. and Ellis, G. (1973) *The Large Scale Structure of Spacetime*, Cambridge University Press.
Hebey, E., Pacard, F., and Pollack, D. 2008, to appear.
Heckmann, O. and Shucking, E. (1962) in *Gravitation, an Introduction to Current Research* (ed. L. Witten), Wiley.
Heusler, M. (1996) *Black Hole Uniqueness Theorems*, Cambridge University Press.
Hoffman, F. and Teller, E. (1950), *Phys. Rev. Ser. 2*, **80**, 692–703.
Hollands, S., Ishibashi, A., and Wald, R. (2007) *Commun. Math. Phys.*, **271**(3), 699–722.
Holst, M., Nagy, G. Tsogtgerel, G. (2008) gr-gc 0802–1031.
Hu Hesheng (1974) *J. Fudan University*, **2**, 92–8.
Huisken G. and Ilmanen T. (2001) J. diff. geona. 59, 353–437.
I-Shih Liu, I., Müller, T. and Ruggeri, T. (1986) *Ann. Phys.*, **169**(1), 191–219.

Ionescu, D. and Klainermanz, S. (2007) gr.qc 0711.0040v1
Isaacson, R. (1968) *Phys. Rev.*, **166**, 1263–80.
Isenberg, J. (1987) *Class. Quant. Grav.*, **12**, 12–49.
Isenberg, J. (1995) *Class. Quant. Grav.*, **12**, 22–49.
Isenberg, J. and Moncrief, V. (1982) *Commun. Math. Phys.*, **86**, 485–93.
Isenberg, J. and Moncrief, V. (1990) *Ann. Phys.*, **199**, 84–122.
Isenberg, J. and Moncrief, V. (2002) *Class. Quant. Grav.*, **19**, 5361–86.
Isenberg, J. and Park (1996) gr.qc 9610027 Vl.
Isenberg, J., Maxwell, D., and Pollack, D. (2005) *Adv. Theor. Math. Phys.*, **9**, 129–72.
Isenberg, J., Mazzeo, R., and Pollack, D. (2002) *Commun. Math. Phys.*, **231**, 529–68.
Isenberg, J., Mazzeo, R., and Pollack, D. (2003) *Ann. H. Poincaré*, **4**, 369–83.
Israel, W. (1976) *Ann. Phys.*, **100**.
Kazdan, J. (1985) Prescribing the curvature of a Riemannian manifold. *Regional Conference Series in Maths*, No. 57, AMS, Providence RI.
Kerner, R. (1968) Ann. H. Poincaré.
Khalatnikov, I. M., Lifshitz, E. M., Khanin, K. M., Shchur, L. N., and Sinai, Ya. G. (1985) *J. Stat. Phys.*, **38**, 97–114.
Kichenassamy, S. and Rendall, A. (1998) *Class. Quant. Grav.*, **15**(5), 1339–55.
Kidder, L., Scheel, M., and Teukolsky, S. (2001) *Phys. Rev. D*, **64**(6).
Klainerman, S. (1986) *Lect. Appl. Math.*, **23**, 293–326.
Klainerman, S. and Machedon, M. (1995) *M. Ann. Math.*, **142**, 39–119.
Klainerman, S. and Nicolo, F. (2002) *The Evolution Problem in General Relativity*, Birkhäuser.
Klainerman, S. and Rodnianski, I. (2002) *C. R. Acad. Sci. Paris*, Serie I, **334**, 125–30.
Klainerman, S. and Rodnianski, I. (2006) arXiv:math.0603009v1.
Klainerman, S. and Rodnianski, I. (2007) *J. Hyp. Diff. Eq.*, **4**(3), 401–33.
Klainerman, S. and Rodnianski, I. (2008) as xiv. math Apl. 709–0801.
Kobayashi, S. and Nomizu, K. (1963) *Differential Geometry*, Vol. 1, Chapter 5, Prentice Hall.
Lacaze, J. (1982) *C. R. Acad. Sci. Paris.*
Leray, J. (1953) *Hyperbolic Differential Equations*, Lecture notes, IAS Princeton.
Leray, J. and Ohya, Y. (1968) *Math. Annalen.*, **162**, 228–36.
Lichnerowicz, A. (1939) *Problèmes Globaux en Mécanique Relativiste*, Herman.
Lichnerowicz, A. (1944) *J. Math. Pures Appl.*, **23**, 39–63.
Lichnerowicz, A. (1955) *Théories Relativistes de la Gravitation et de l'éléctromagnétisme*, Masson.
Lichnerowicz, A. (1963) *C. R. Acad. Sci. Paris*, **7**.
Lichnerowicz, A. (1964) *Ann. Sci. IHES*, No. 1.
Lichnerowicz, A. (1967) *Relativistic Hydrodynamics and Magnetohydrodynamics*, Benjamin, New York.
Lindblad, H. and Rodnianski, I. (2005) *Commun. Math. Phys.*, **246**(1), 43–110 and arXiv:math/0411109v1.
Lindblom, L., Matthews, K., Rinne, O., and Scheel, M. (2008) gr-qc 0711-2084.
Loizelet, J. (2005) *C. R. Acad. Sci. Paris*, Serie I, **340**.
Müller, I. (1966) *PhD Thesis*, Aachen.

Müller, I. and Ruggeri, T. (1998) *Rational Extended Thermodynamics*, Springer Tracts in Natural Philosophy, 37 (2nd Edition), Springer-Verlag, New York
Müller zum Hagen, H., Seifert, H. J., and Yodzis, P. (1973) *Comm. Math. Phys.*, **34**, 135–48).
Majumdar, S. D. (1947) *Phys. Rev.*, **72**, 390–8.
Marle, C. M. (1969) *Ann. Inst. Poincaré* **2**, 67–126 and 127–94.
Mars, M. (2000) *Class. Quant. Grav.*, **17**, 335–73.
Maxwell, D. (2003) gr-qc 0307117.
Maxwell, D. (2005) *J. Hyp. Diff. Eq.*, **2**, 521–46.
Maxwell, D. (2006) *J. Reine Angew. Math.*, 1–29.
Milnor, J. (1965) *Topology from the Differentiable Viewpoint*, University Press of Virginia, Charlotteville.
Moncrief, V. (1975) *J. Math. Phys.*, **16**, 493–8 and **17**, 1893–902.
Moncrief, V. (1981) *Ann. Phys.*, **132**, 87–107.
Moncrief, V. (1986) *Ann. Phys.*, **167**(1), 118–42.
Moncrief, V. (1990) *Class. Quant. Grav.*, **7**, 320–52.
Moncrief, V. (2005) Isaac Newton Institute Preprints.
Moncrief, V. and Isenberg, J. (1983) *Commun. Math. Phys.*, **89**, 387–413.
Moncrief, V. and Isenberg, J. (2008) *Class. Quant. Grav.*, **25**, 19015.
Morawetz, C. (1961) *Commun. Pures. Appl. Math.*, **14**, 561–8.
Muller, S. and Struwe, M. (1996) *Top. Meth. Non-Lin. Anal.*, **7**(2), 245–61.
Newman, R. (1986) *Class. Quantum Grav.*, **3**, 527–39.
Nomizu, K. and Ozeki, H. (1961) *Proc. Am. Math. Soc.*, **12**, 889–92.
O'Murchada, N. and York, J. (1973) *J. Math.*, **14**, 1551–7.
O'Murchada, N. and York, J. (1974) *Phys. Rev. D*, **10**, 428–46.
Ohanian, H. and Ruffini, R. (1994) *Gravitation and Spacetime*, 2nd edn, Ch. 4, Norton.
Oppenheimer, J. R. and Snyder, H. (1939) *Phys. Rev.* **56**, 455.
Papapetrou, A. (1947) *Proc. R. Irish Acad.*, **A51**, 191–204.
Papapetrou, A. and Hamoui, A. (1967) *Ann. Inst. Poincaré A*, **VI**, 343–64.
Penrose, R. (1965) *Phys. Rev. Lett.*, **14**, 57.
Penrose, R. (1968) in *Batelle Rencontres* (eds. C. DeWitt and J. A. Wheeler), Benjamin, New York.
Penrose, R. (1979) In *An Einstein Centenary Survey* (eds. S. Hawking and W. Israel), Cambridge University Press.
Pichon, G. (1965) *Ann. Inst. Henri Poincaré A*, **2**, 21–85.
Pretorius, F. (2005) *Phys. Rev. Lett.*, **95**, 121.
Pretorius, F. (2005) *Class. Quant. Grav.*, **22** 425–52.
Protter, M. H. and Weinberger, H. F. (1967) *Maximum Principles in Differential Equations*, Prentice Hall, p. 64.
Rein, D. and Rendall, A. (1992) *Commun. Math. Phys.*, **150**, 561–83.
Rendall, A. (1992) *J. Math. Phys.*, **33**, 1047–53.
Rendall, A. (2000) *Class. Quant. Grav.*, **17**, 3305–16.
Rendall, A. (2004) *Class. Quant. Grav.*, **21**, 2445–54.
Ringström, H. (2001) *Ann. H. Poincaré*, **2**, 405–500.
Ringström, H. (2004) *Class. Quant. Grav.*, **21**(3), S305–S322.
Ringström, H. (2004) *Math. Proc. Cambridge Philos. Soc.*, **136**(2), 485–512.
Ringström, H. (2004) *Commun. Pure Appl. Math.*, **57**(5), 657–703.

Ringström, H. (2006) *Commun. Pure Appl. Math.*, **59**, 977–1041.
Ringström, H. (2008) in *Invent. Math.*, **173**, 123–208.
Rionero, S. and Ruggeri, T. (1992) (eds.) *Waves and Stability in Continuous Media*, pp. 54–69, World Scientific, Singapore.
Ruback, P. (1988) *Class. Quant. Grav.*, **5**, L155–9.
Ruggeri, T. (1987) *Corso CIME Noto*, Lecture Notes in Mathematics 1385 (eds. A. Anile and Y. Choquet-Bruhat), pp. 269–77, Springer-Verlag, Berlin.
Ruggeri, T. (2004) *Il Nuovo Cimento B*, **119**(7–9), 809–21 (*Lecture Notes of the International Conference in honour of Y. Choquet-Bruhat: Analysis, Manifolds and Geometric Structures in Physics*, eds. G. Ferrarese and T. Ruggeri).
Ruggeri, T. and Serre, D. (2003) *Q. Appl. Math.*, **62**(1), 163.
Ruggeri, T. and Strumia, A. (1981) *Ann. Inst. H. Poincaré*, **34A**, 65–84.
Ryan, M. and Shepley, L. (1975) *Homogeneous Relativistic Cosmologies*, Princeton University Press.
Schneider, P., Ehlers, J., and Falco, E. E. (1992) *Gravitational Lenses*. Springer-Verlag, Berlin.
Schoen, R. (1984) *J. Diff. Geom.*, **20**, 479–95.
Segal, I. E. (1963) *Ann. Math.*, **78**(2), 339–64.
Segal, I. E. (1976) Mathematical cosmology, Acad. Press.
Shizuta, Y. and Kawashima, S. (1985) *Hokkaido Math. J.*, **14**, 249–75.
Simon, W. (1992) *Class. Quant. Grav.*, **9**, 241–56.
Sobolev, S. (1936) *Math. Sbornik*, **1**(N.S.), 39–71.
Sogge, C. (1995) *Lectures on Non-Linear Wave Equations*, International Press.
Sudarsky, D. and Wald, R. (1993) *Phys. Rev. D*, **12**(3), 5209–13.
Taub, A. H. (1957) *Illinois J. Math.*, **1**, 370–88.
Taub, A. H. (1959) *Arch. Rat. Mech. An.*, **3**, 312–29.
Taub, A. H. (2004) *General Relativity and Gravitation*, **36**(12), 1699–719.
Teitelboim (1983) (*Phys. Rev. D*, **28**, 297–311.
Tolman, R. (1934) *Proc. Natl. Acad. Sci. USA*, **20**, 169–76.
Trautman, A. (1964) In *Brandeis Summer Institute in Theoretical Physics*, Prentice Hall.
Trudinger, N. (1968) *Ann. Norm. Sup. Pisa*, **22**, 265–74.
Vaillant, J. (1968) *J. Math. Pures Appl.*, **47**, 1–40.
Vaillant-Simon, A. (1969) *J. Math. Pures Appl.*, **48**.
Wald, R. (1984) *General Relativity*, University of Chicago Press.
Weaver, M., Isenberg, J., and Berger, B. K. (1998) *Phys. Rev. Lett.*, **80**, 2984.
Whitney, H. Integration theory Princeton u. press.
Wolf, J. *Spaces of Constant Curvature*, Chapter 3, Publish or Perish Inc., Houston TX.
Yong, W. A. (2004) *Arch. Rat. Mech. Anal.*, **172**, 247.
York, J. W. (1972) *Phys. Rev. Lett.*, **28**, 1082.
York, J. W. (1974) *Ann. Inst. Henri Poincaré*, **21**, 319–32.

This page intentionally left blank

INDEX

Abstract index notation, 4
Acausal, 376
Achronal, 376, 393
ADM equations, 156
Alfvén cone, 295
Alfvén waves, 295
Algebraic identity, 13
Anti de Sitter spacetime, 120
Apparent horizon, 417, 419
Aspherical, 138
Asymptotically Euclidean manifold, 221
  with boundary, 460
  with non-compact interior, 460
Asymptotically strongly predictable, 419
Asymptotically Velocity Term Dominated (AVTD), 403
Atlas, 1
Axisymmetric, 467

Bach tensor, 498
Back reaction, 358
Barotropic fluids, 54
Beginning and age of the universe, 122
Behaviour with increasing t, 123
Belinskii, Khalatnikov, and Lifshitz (BKL) conjecture
  in Iwasawa variables, 440
Bel–Robinson energy, 256, 530
Bel tensor, 254, 255, 530
Bianchi classification, 127
Bianchi identity, 13
Bianchi models, 126
Birkhoff theorem, 75
Black hole region β, 419
Bondi mass, 427
Bounded curvature conjecture, 615

Canonical choice, 575
Carter–Robinson theorem, 475
Cattaneo–Ferrarese frame, 285
Cauchy adapted frame, 146
Cauchy horizon, 394
Cauchy regular, 622
Cauchy surface, 371, 394
Causal future, 376
Causal structure, 374, 377
Causal subset, 170, 578
Causal vector, 9
Characteristic cone, 618
Characteristic determinant, 622
Characteristic equation, 344
Characteristic tangent vectors, 618
Chart, 1
Christoffel symbols, 11
Chronological past, 376
Circular spacetimes, 468
Codazzi tensors, 529
Codifferential operator, 11
Comoving/proper frame, 34
Compact towards the past, 620
Complete Riemannian manifold, 372
Cone property, 535
Conformal, 660
Conformal Killing fields, 203
Conformal thin sandwich, 198
Connection, 656
Connection coefficients, 10, 656
Conservation laws, 146
Conservation of energy, 581
Continuity equation, 260
Contracted Bianchi identity, 14
Contracted product, 5
Contravariant tensor, 5
Cosmic censorship conjecture, 103
Cosmic time, 107
Cosmological constant, 192
Cosmological principle, 107
Cotangent space, 3
Coton tensor, 498, 646
Covariant derivative, 9, 10
Covariant differential, 50
Covariant tensor, 5
Covariant vectors, 3
Cronström gauge, 615
Curvature, 656
Curvature tensor, 13
Curve, 375, 389

Damped wave coordinates, 162
Deceleration parameter, 117
Degenerate horizon, 477
Degree of symmetry, 137
Density parameter, 125
Derivation, 4
Derivation formula, 62
de Sitter spacetime, 118
Development, 156, 400

Differentiable manifold, 1
Distortion of signals, 347
Distribution function, 302
Domain of dependence, 394
Domain of dependence property, 620
Domain of outer communications, 477
Dominant energy condition, 51, 267
Double null foliation, 504
Dual frames, 3
Dust, 54

Eikonal, 270, 618
Einstein–de Sitter universe, 123
Einstein equations, 43
  in vacuum, 43
Einsteinian development, 156
Einstein static universe, 115
Einstein tensor, 14, 145
Einstein–Yang–Mills solitons, 466
Electromagnetic field, 45
Ellipticity constant, 242
End, 537
End point, 375
Energy equality, 577, 626
Energy estimate, 626
Energy functional, 626
Energy inequality, 318, 626
Energy momentum vector, 31, 267
Equation of state, 54
Ergosphere, 473
Euclidean at infinity, 537
Euclidean space, 537
Euclidean space form, 138
Euler equations, 266
Event horizon, 92, 419, 652
Exceptional characteristic manifolds, 346
Exceptionality condition, 65, 357
Expansion, 409
Exponential mapping, 383
Extendible, 375
Exterior derivative, 6
Exterior global smallness condition, 504
Exterior p-form, 6
Extrinsic curvature, 147

Family of timelike geodesics
    parametrized, 410
Final Bondi mass, 433
First moment, 305
First-order symmetric hyperbolic (FOSH)
    systems, 244, 617, 628
First variation, 17
Fisher–Marsden criterion, 185
Flat space, 8
Flat type, 138
Fluid source, 259

Frame, 4
Friedman equation, 114
Friedmann–Lemaître models, 122
Friedrich conformal system, 490
Fundamental observers, 107
Fundamental subsets, 378
Future, 9
Future completeness, 404

Galileo group, 20
Gauss formula, 456
Generalized null condition, 359
Genus, 139
Geodesic, 12
Geodesic congruence, 410
Geodesic coordinates, 383
Geodesic deviation, 15, 411
Geodesic motion, 260
Gevrey classes, 624
Gevrey index, 624
Global hyperbolicity, 371, 620
Gravitational collapse, 92
Gravitational lens, 39
Gravitational wave, 60
Gromwall lemma, 580

Half polarization, 450
Hall effect, 46
Hamiltonian constraint, 179, 194
Harmonically flat at infinity, 420
Harmonic map, 50
Harmonic/wave coordinates, 61, 158
Hausdorf manifolds, 2
H-energy, 316
Higher moments, 306
Higher order covariant derivatives, 50
High-frequency wave, 62, 342
Hodge adjoint, 11
Homogeneous, 109
Homogeneous cosmology, 125
Hopf–Rinow theorem, 372
Hubble constant, 116
Hubble law, 117
Hyperbolic at, 618
Hyperbolic polynomial, 618
Hyperbolic space form, 138
Hyperbolic system, 617
Hyperbolic type, 138
Hyperboliques non strictes, 624

Incompressible fluid, 273
Inertial frames, 20
Infinite conductivity, 294, 297
Initial data set, 156
Irrotational, 408

Isometry, 7, 653
Israel theorem, 475
Iwasawa decomposition, 438

Jordan decomposition, 635

Kantowski-Sachs, 136
Kasner exponents, 130
Kasner spacetimes, 128
Killing horizon, 477
Kirchhoff–Sobolev formula, 615
Komar mass, 460
Kretschmann scalar, 93, 472

Lanczos identity, 254
Landau–Lifshitz pseudo tensor, 464
Lapse, 146
Lemaître, 124
Length, 7
Leray-hyperbolic system, 617, 622
Leray–Ohya hyperbolic, 624
Lichnerowicz equation, 189
Lie derivative, 6
Linearization, 17
Linearization stable, 181
Liouville–Vlasov equation, 307, 310
Local coordinates, 1
Local energy inequality, 579
Local $H_s(V)$ metrics, 56
Locally isometric, 8
Locally rotationally symmetric (LRS), 126, 135
Lorentz condition, 47
Lorentz contraction, 26
Lorentz dilatation, 26
Lorentz force, 23
Lorentz group, 24
Lorentzian manifold, 374
Lorentzian metric, 8, 143

Magnetoacoustic wave fronts, 295
Magnetohydrodynamics, 46
Majumdar–Papapetrou solutions, 476
Manifold, 55
Mass hyperboloid, 303
Maximal, 400
Maxwell–Jüttner distributions, 334
Maxwell's equations, 45
Maxwell tensor, 263
Mean curvature, 147
Michelson–Morley experiments, 29
Milne universe, 124
Momentum constraint, 179, 194
Moncrief theorem, 184
Mössbauer resonant effect, 41
Multiplication property, 536

Naked singularity, 421
Natural coframe, 3
Natural frame, 3
Navier–Stokes equations, 298
No naked singularity conjecture, 422
Normal geodesic coordinates, 384
Normal neighbourhood, 384
Null average *ansatz*, 63
Null condition, 358
Null cone, 432
Null dust, 260
Null mean curvatures, 415
Null structure, 502
Null vector, 9

Optical function, 502
Orbit, 653
Orientable, 2
Orthochronous, 24
Orthonormal Lorentzian frames, 8

Paracompact, 371
Parallel transport, 12
Past, 9
Path/parametrized curve, 375
Penrose confomorphism, 647
Penrose process, 474
Perfect fluid, 35, 54
Phase space, 302
Piecewise differentiable, 375
Poincaré group, 23
Poincaré inequality, 541
Polarization condition, 60, 64, 69, 344, 360
Polarization operator, 365
Polarization vector, 350
Polarized case, 443
Polarized exceptionality condition, 361
Polarized null condition, 361
Principle of general covariance, 39
Progressive wave, 62, 342
Propagation equation, 361
Propagation speed, 271
Proper frame, 28, 259, 265
Proper time, 26, 39
Pseudo norm, 7
Pseudo-Riemannian metric, 6
Pullback, 5

Quasidiagonal by blocks, 623
Quasidiagonalisation, 622

Rankine–Hugoniot equations, 274

Raychauduri equation, 411
Red shift, 41
Red shift parameter, 115
Regular, 432
Regularly asymptotically, 419
Regularly sliced, 397, 585
Regular vector field, 372
Reissner–Nordstörm solution, 103
Representative, 3
Ricci identity, 13, 50
Ricci tensor, 14, 145
Riemann curvature tensor, 145
Riemannian connection, 10, 144
Riemannian metric, 8
Rigid, 529
Robertson–Walker spacetimes, 111
Rotational, 408

Scalar curvature, 14, 145
Scalar multiplet equations, 362
Scalar multiplet Φ, 362
Schwarzschild radius, 78
Schwarzschild spacetime, 78
Second fundamental form, 148
Second moment, 305
Second variation, 17, 18
Sectional curvature, 108
Segal cosmos, 115
Shear, 409
Shift, 146
  σ-constant, 209
Simply transitively, 653
Sliced spacetime, 146
Sobolev regular, 536, 594
Soliton, 458
Space of constant Riemannian curvature, 109
Spacetime, 39, 374
Special Lorentz, 25
Spherical space form, 138
Spherical type, 138
Spin-2 field, 488
stability of circular orbits, 83
Stably causal, 392
Standard cosmological models, 107
Standard null forms, 488
Static, 451
Stationary asymptotically Euclidean spacetimes, 457
Stationary axisymmetric, 467
Stationary spacetimes, 451
Stiff fluid, 273
Stokes regular, 576
Stress energy tensor, 51
Strong cosmic censorship conjecture, 401
Strong energy condition, 51, 266, 411
Strong equivalence principle, 44
Strong hyperbolicity, 627
Strongly asymptotically Euclidean, 498
Strongly causal, 392
Structure coefficients, 143, 655
Structure constants, 126
Subcharacteristic, 347
Subsolution, 563
Supersolution, 563
Surface gravity, 477
SY class, 225
Symmetrizable hyperbolic systems, 627

Tangent space, 3
Tangent vector, 3
Taub conserved quantity, 184
Taub-NUT spacetime, 136
Taub spacetime, 135
Temporal gauge, 47
Tensor fields, 5
Tensor product, 5
Thurston classification conjecture, 139
Time adapted frame, 284
Time delay, 42
Time function, 392, 397
Timelike, 9
Time orientable, 9, 374
Time oriented, 619
Time oriented Lorentzian manifold, 374
Time symmetric, 417
TIP (Terminal indecomposable past set), 423
Tolman–Oppenheimer–Volkov equations, 90
Topological manifold, 1
Total Bondi mass, 427
Trapped surfaces, 414
TT (Transverse, Traceless) tensors, 188
Twin paradox, 27
Twist potential, 516
Two timing method, 62

Uniqueness and causality theorem, 580
Unit velocity, 30

Vlasov–Maxwell–Einstein system, 327
Vlasov models, 307
Volume element, 7
Volume form, 7
Vorticity, 409

Wave front, 271
Wave map, 50
Weak cosmic censorship conjecture, 403, 422, 423
Weak energy condition, 51, 266
Weak equivalence principle, 39

Weak progressive gravitational wave, 62
Weighted Sobolev space, 221, 537
Weyl tensor, 645
Wormhole, 234

Yamabe classes, 205
Yamabe equation, 206
Yamabe number, 205
Yamabe types, 208
Yang–Mills equations, 362
Yang–Mills field, 192, 361
Yang–Mills fluids, 297
York-scaled, 189
York splitting, 197