An introduction to the Cauchy problem for the Einstein equations

Roscoff, November 2010

all comments on misprints and suggestions for improvements welcome

Piotr T. Chruściel
University of Vienna
homepage.univie.ac.at/piotr.chrusciel

March 18, 2011

Contents

Part I

The Einstein equations

Chapter 1

The Einstein equations

1.1 The nature of the Einstein equations

The *vacuum Einstein equations with cosmological constant* Λ read

$$G_{\alpha\beta} + \Lambda g_{\alpha\beta} = 0 \,, \tag{1.1.1}$$

where $G_{\alpha\beta}$ is the Einstein tensor,

$$G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} \,, \tag{1.1.2}$$

while $R_{\alpha\beta}$ is the Ricci tensor and R the scalar curvature. We will sometimes refer to those equations as *the vacuum Einstein equations*, regardless of whether or not the cosmological constant vanishes. Taking the trace of (1.1.1) one obtains

$$R = \frac{2(n+1)}{n-1}\Lambda \,, \tag{1.1.3}$$

where, as elsewhere, $n+1$ is the dimension of space-time. This leads to the following equivalent version of (1.1.1):

$$\mathrm{Ric} = \frac{2\Lambda}{n-1} g \,. \tag{1.1.4}$$

Thus the Ricci tensor of the metric is proportional to the metric. Pseudo-Lorentzian manifolds the metric of which satisfies Equation (1.1.4) are called *Einstein manifolds* in the mathematical literature; see, e.g., [25].

Given a manifold $\mathscr{M}$, Equation (1.1.1) or, equivalently, Equation (1.1.4) forms a system of partial differential equations for the metric. Indeed, recall that

$$\Gamma^\alpha{}_{\beta\gamma} = \tfrac{1}{2} g^{\alpha\sigma}(\partial_\beta g_{\sigma\gamma} + \partial_\gamma g_{\sigma\beta} - \partial_\sigma g_{\beta\gamma}) \,, \tag{1.1.5}$$

$$R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha{}_{\beta\delta} - \partial_\delta \Gamma^\alpha{}_{\beta\gamma} + \Gamma^\alpha{}_{\sigma\gamma}\Gamma^\sigma{}_{\beta\delta} - \Gamma^\alpha{}_{\sigma\delta}\Gamma^\sigma{}_{\beta\gamma} \,, \tag{1.1.6}$$

$$R_{\alpha\beta} = R^\gamma{}_{\alpha\gamma\beta} \,. \tag{1.1.7}$$

We see that the Ricci tensor is an object built out of the Christoffel symbols and their first derivatives, while the Christoffel symbols are built out of the metric and its first derivatives. These equations further show that the Ricci tensor is linear in the second derivatives of the metric, with coefficients which are rational functions of the $g_{\alpha\beta}$'s, and quadratic in the first derivatives of g, again with coefficients rational in g. Equations linear in the highest order derivatives are called *quasi-linear*, hence the vacuum Einstein equations constitute a second order system of quasi-linear partial differential equations for the metric g.

In the discussion above we have assumed that the manifold $\mathscr{M}$ has been given. Such a point of view might seem to be too restrictive, and sometimes it is argued that the Einstein equations should be interpreted as equations both for the metric and the manifold. The sense of such a statement is far from being clear, one possibility of understanding that is that the manifold arises as a result of the evolution of the metric g. We are going to discuss in detail the evolution point of view below, let us, however, anticipate and mention the following: there exists a natural class of space-times, called *maximal globally hyperbolic*, which are obtained by the vacuum evolution of initial data, and which have topology $\mathbb{R}\times\mathscr{S}$, where $\mathscr{S}$ is the n-dimensional manifold on which the initial data have been prescribed. Thus, these space-times have topology and differentiable structure which are determined by the initial data. It turns out that the space-times so constructed are sometimes *extendible.* Now, there do not seem to exist conditions which would guarantee uniqueness of extensions of the maximal globally hyperbolic solutions, while examples of non-unique extensions are known. Therefore it does not seem useful to consider the Einstein equations as equations determining the manifold beyond the maximal globally hyperbolic region. We conclude that in the evolutionary point of view the manifold can be also thought as being given *a priori*, namely $\mathscr{M}=\mathbb{R}\times\mathscr{S}$. We stress, however, that there is no natural time coordinate which can always be constructed by evolutionary methods and which leads to the decomposition $\mathscr{M}=\mathbb{R}\times\mathscr{S}$.

Now, there exist standard classes of partial differential equations which are known to have good properties. They are determined by looking at the algebraic properties of those terms in the equations which contain derivatives of highest order, in our case of order two. Inspection of (1.1.1) shows (cf., e.g., [75]) that this equation does not fall in any of the standard classes, such as hyperbolic, parabolic, or elliptic. In retrospect this is not surprising, because equations in those classes typically lead to unique solutions. On the other hand, given any solution g of the Einstein equations (1.1.4) and any diffeomorphism Φ, the pull-back metric Φ^*g is also a solution of (1.1.4), so whatever uniqueness there might be will hold only *up to diffeomorphisms.* An alternative way of describing this, often found in the physics literature, is the following: suppose that we have a matrix $g_{\mu\nu}(x)$ of functions satisfying (1.1.1) in some coordinate system x^μ. If we perform a coordinate change $x^\mu\to y^\alpha(x^\mu)$, then the matrix of functions $\bar g_{\alpha\beta}(y)$ defined as

$$g_{\mu\nu}(x)\to\bar g_{\alpha\beta}(y)=g_{\mu\nu}(x(y))\frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta}\tag{1.1.8}$$

will also solve (1.1.1), if the x-derivatives there are replaced by y-derivatives. This property is known under the name of *diffeomorphism invariance*, or *co-*

ordinate invariance, of the Einstein equations. Physicists say that "the diffeomorphism group is the gauge group of Einstein's theory of gravitation".

Somewhat surprisingly, Choquet-Bruhat [68] proved in 1952 that there exists a set of *hyperbolic* equations underlying (**??**). This proceeds by the introduction of so-called "harmonic coordinates", to which we turn our attention in the next section. Before doing that, let us pass to the derivation of a somewhat more explicit *and useful* form of the Einstein equations. In index notation, the definition of the Riemann tensor takes the form

$$\nabla_\mu \nabla_\nu X^\alpha - \nabla_\nu \nabla_\mu X^\alpha = R^\alpha{}_{\beta\mu\nu} X^\beta \; . \tag{1.1.9}$$

A contraction over α and μ gives

$$\nabla_\alpha \nabla_\nu X^\alpha - \nabla_\nu \nabla_\alpha X^\alpha = R_{\beta\nu} X^\beta \; . \tag{1.1.10}$$

Suppose that X is the gradient of a function ϕ, $X = \nabla\phi$, then we have

$$\nabla_\alpha X^\beta = \nabla_\alpha \nabla^\beta \phi = \nabla^\beta \nabla_\alpha \phi \; ,$$

because of the symmetry of second partial derivatives. Further

$$\nabla_\alpha X^\alpha = \Box_g \phi \; ,$$

where we use the symbol

$$\Box_k \equiv \nabla_\mu \nabla^\mu$$

to denote the wave operator associated with a Lorentzian metric k; *e.g.*, for a scalar field we have

$$\Box_g \phi \equiv \nabla_\mu \nabla^\mu \phi = \frac{1}{\sqrt{-\det g_{\alpha\beta}}} \partial_\mu (\sqrt{-\det g_{\rho\sigma}} g^{\mu\nu} \partial_\nu \phi) \; . \tag{1.1.11}$$

For gradient vector fields (1.1.10) can be rewritten as

$$\nabla_\alpha \nabla^\alpha \nabla_\nu \phi - \nabla_\nu \nabla_\alpha \nabla^\alpha \phi = R_{\beta\nu} \nabla^\beta \phi \; ,$$

or, equivalently,

$$\Box_g d\phi - d(\Box_g \phi) = \mathrm{Ric}(\nabla\phi, \cdot) \; , \tag{1.1.12}$$

where d denotes exterior differentiation. Consider Equation (1.1.12) with ϕ replaced by y^A, where y^A is any collection of functions,

$$\Box_g dy^A = d\lambda^A + \mathrm{Ric}(\nabla y^A, \cdot) \; , \tag{1.1.13}$$

$$\lambda^A \equiv \Box_g y^A \; . \tag{1.1.14}$$

Set

$$g^{AB} \equiv g(dy^A, dy^B) \; ; \tag{1.1.15}$$

this is consistent with the usual notation for the inverse metric except that we haven't assumed (yet) that the y^A's form a coordinate system. (Furthermore, for simplicity we write g instead of $g^\flat$ for the metric on T^*M.) By the chain rule we have

$$\begin{aligned}
\Box_g g^{AB} &= \nabla_\mu \nabla^\mu (g(dy^A, dy^B)) \\
&= \nabla_\mu (g(\nabla^\mu dy^A, dy^B) + g(dy^A, \nabla^\mu dy^B)) \\
&= g(\Box_g dy^A, dy^B) + g(dy^A, \Box_g dy^B) + 2g(\nabla_\mu dy^A, \nabla^\mu dy^B) \\
&= g(d\lambda^A, dy^B) + g(dy^A, d\lambda^B) + 2g(\nabla_\mu dy^A, \nabla^\mu dy^B) \\
&\quad +2\mathrm{Ric}(\nabla y^A, \nabla y^B)\ . \qquad (1.1.16)
\end{aligned}$$

Let us *suppose* that the functions y^A solve the homogeneous wave equation:

$$\lambda^A = \Box_g y^A = 0\ . \qquad (1.1.17)$$

The Einstein equation (1.1.4) inserted in (1.1.16) implies then

$$\begin{aligned}
E^{AB} &\equiv \Box_g g^{AB} - 2g(\nabla_\mu dy^A, \nabla^\mu dy^B) - \frac{4\Lambda}{n-1} g^{AB} \qquad (1.1.18a) \\
&= 0\ . \qquad (1.1.18b)
\end{aligned}$$

Now,

$$\begin{aligned}
\nabla_\mu (dy^A) &= \nabla_\mu (\partial_\nu y^A\, dx^\nu) \\
&= (\partial_\mu \partial_\nu y^A - \Gamma^\sigma{}_{\mu\nu} \partial_\sigma y^A) dx^\nu\ . \qquad (1.1.19)
\end{aligned}$$

Suppose that the $d\phi^A$'s are linearly independent and form a basis of $T^*\mathscr{M}$, then (1.1.18b) is *equivalent* to the vacuum Einstein equation. Further we can choose the y^A's as coordinates, at least on some open subset of $\mathscr{M}$; in this case we have

$$\partial_A y^B = \delta^B_A\ , \qquad \partial_A \partial_C y^B = 0\ ,$$

so that (1.1.19) reads

$$\nabla_B dy^A = -\Gamma^A{}_{BC} dy^C\ .$$

This, together with (1.1.18b), leads to

$$\Box_g g^{AB} - 2g^{CD} g^{EF} \Gamma^A_{CE} \Gamma^B_{DF} - \frac{4\Lambda}{n-1} g^{AB} = 0\ . \qquad (1.1.20)$$

Here the Γ^A_{BC}'s should be calculated in terms of the g_{AB}'s and their derivatives as in the usual equation for the Christoffel symbols, and the wave operator $\Box_g$ is understood as acting on scalars. We have thus shown that *in "wave coordinates", as defined by the condition* $\lambda^A = 0$, *the Einstein equation forms a second-order quasi-linear wave-type system of equations* (1.1.20) for the metric functions g^{AB}. This gives a strong hint that the Einstein equations possess a *hyperbolic*, evolutionary character; this fact will be fully justified in what follows.

Another completely explicit form *without imposing any coordinate conditions*, which is not very enlightening, and fortunately almost never needed, reads

$$\begin{aligned} R_{\nu\rho}[g] \quad = \quad & \frac{1}{2}\left\{\frac{\partial}{\partial x^\delta}\left(g^{\delta\eta}\left[-\frac{\partial g_{\rho\nu}}{\partial x^\eta}+\frac{\partial g_{\nu\eta}}{\partial x^\rho}+\frac{\partial g_{\rho\eta}}{\partial x^\nu}\right]\right)-\frac{\partial}{\partial x^\rho}\left(g^{\delta\eta}\frac{\partial g_{\delta\eta}}{\partial x^\nu}\right)\right\} \\ & +\frac{1}{4}\left\{g^{\lambda\pi}\left(\frac{\partial g_{\delta\pi}}{\partial x^\lambda}+\frac{\partial g_{\lambda\pi}}{\partial x^\delta}-\frac{\partial g_{\lambda\delta}}{\partial x^\pi}\right)g^{\delta\eta}\left(\frac{\partial g_{\nu\eta}}{\partial x^\rho}+\frac{\partial g_{\rho\eta}}{\partial x^\nu}-\frac{\partial g_{\rho\nu}}{\partial x^\eta}\right)\right. \\ & \left.-g^{\lambda\eta}\left(\frac{\partial g_{\delta\eta}}{\partial x^\rho}+\frac{\partial g_{\rho\eta}}{\partial x^\delta}-\frac{\partial g_{\rho\delta}}{\partial x^\eta}\right)g^{\delta\pi}\left(\frac{\partial g_{\nu\pi}}{\partial x^\lambda}+\frac{\partial g_{\lambda\pi}}{\partial x^\nu}-\frac{\partial g_{\lambda\nu}}{\partial x^\pi}\right)\right\} . \end{aligned} \tag{1.1.21}$$

It should be kept in mind that the coefficients $g^{\delta\eta}$ of the matrix $(g^{\delta\eta})$ inverse to $(g_{\mu\nu})$ take the form $g^{\delta\eta} = (\det(g_{\mu\nu}))^{-1} p^{\delta\eta}$, with $p^{\delta\eta}$'s being homogeneous polynomials, of degree one less than the dimension of the manifold, in the $g_{\mu\nu}$'s.

It turns out that (1.1.18b) allows one also to *construct* solutions of Einstein equations [68], this will be done in the following sections.

Before analyzing the existence question, it is natural to ask the following: given a solution of the Einstein equations, can one always find local coordinate systems y^A satisfying the wave condition (1.1.17)? The answer is yes, the standard way of obtaining such functions proceeds as follows: Let $\mathscr{S}$ be any spacelike hypersurface in $\mathscr{M}$; by definition, the restriction of the metric g to $T\mathscr{S}$ is positive non-degenerate. Let $\mathscr{O} \subset \mathscr{S}$ be any open subset of $\mathscr{S}$, and let X be any smooth vector field on $\mathscr{M}$, defined along $\mathscr{O}$, which is transverse to $\mathscr{S}$; by definition, this means that for each $p \in \mathscr{O}$ the tangent space $T_p\mathscr{M}$ is the direct sum of $T_p\mathscr{S}$ and of the linear space $\mathbb{R}X(p)$ spanned by $X(p)$. (Any timelike vector X would do — *e.g.*, the unit normal to $\mathscr{S}$ — but transversality is sufficient for our purposes here.) The following result is well known, though difficult to locate in the literature:

THEOREM 1.1.1 *For any smooth functions f, g on $\mathscr{O} \subset \mathscr{S}$ there exists a unique smooth solution ϕ defined on $\mathscr{D}(\mathscr{O})$ of the problem*

$$\Box_g \phi = 0 , \qquad \phi|_{\mathscr{O}} = f , \quad X(\phi)|_{\mathscr{O}} = g .$$

Once a hypersurface $\mathscr{S}$ has been chosen, *local wave coordinates adapted to $\mathscr{S}$* may be constructed as follows: Let $\mathscr{O}$ be any coordinate patch on $\mathscr{S}$ with coordinate functions x^i, $i = 1, \ldots, n$, and let e^0 be the field of unit future pointing normals to $\mathscr{O}$. On $\mathscr{D}(\mathscr{O})$ define the y^A's to be the unique solutions of the problem

$$\Box_g y^A = 0 ,$$

$$y^0|_{\mathscr{O}} = 0 , \quad e^0(y^0)|_{\mathscr{O}} = 1 , \tag{1.1.22}$$

$$y^i|_{\mathscr{O}} = x^i , \quad e^0(y^i)|_{\mathscr{O}} = 0 , \quad i = 1, \ldots, n . \tag{1.1.23}$$

We note that there is a considerable freedom in the construction of the y^i's (because of the freedom of choice of the x^i's), but the function y^0 is defined uniquely by $\mathscr{S}$. Since the x^i's form a coordinate system on $\mathscr{O}$, a simple application of the implicit function theorem shows that there exists a neighborhood $\mathscr{U} \subset \mathscr{D}(\mathscr{O})$ of $\mathscr{O}$ which is coordinatized by the y^A's.

1.2 Existence local in time and space in wave coordinates

Let us return to (1.1.16). Assume again that the y^A's form a local coordinate system, but do not assume for the moment that the y^A's solve the wave equation. In that case (1.1.16) together with the definition (1.1.18a) of E^{AB} lead to

$$R^{AB} = \frac{1}{2}(E^{AB} - g^{AC}\partial_C\lambda^B - g^{BC}\partial_C\lambda^A) + \frac{2\Lambda}{n-1}g^{AB} . \qquad (1.2.1)$$

For the purpose of the calculations that follow, it turns out to be convenient to treat the index A on the λ's as a vector index, and change the partial derivatives in (1.2.1) to vector-covariant ones:

$$\begin{aligned} & E^{AB} - g^{AC}\partial_C\lambda^B - g^{BC}\partial_C\lambda^A = \\ & \quad \underbrace{E^{AB} + g^{AC}\Gamma^B{}_{CD}\lambda^D + g^{BC}\Gamma^A{}_{CD}\lambda^D}_{=:\hat{E}^{AB}} \\ & \quad -g^{AC}(\partial_C\lambda^B + \Gamma^B{}_{CD}\lambda^D) - g^{BC}(\partial_C\lambda^A + \Gamma^A{}_{CD}\lambda^D) . \end{aligned} \qquad (1.2.2)$$

One can then rewrite (1.2.1) as

$$R^{AB} = \frac{1}{2}(\hat{E}^{AB} - \nabla^A\lambda^B - \nabla^B\lambda^A) + \frac{2\Lambda}{n-1}g^{AB} . \qquad (1.2.3)$$

The idea, due to Yvonne Choquet-Bruhat [68], is to use the hyperbolic character of the equation

$$\hat{E}^{AB} = 0 \qquad (1.2.4)$$

to construct a metric g. If we manage to make sure that λ^A vanishes as well, it will then follow from (1.2.1) then g will also solve the Einstein equation. The following result is again standard:

Theorem 1.2.1 *For any initial data*

$$g^{AB}(y^i, 0) \in H^{k+1} , \quad \partial_0 g^{AB}(y^i, 0) \in H^k , \qquad k > n/2 , \qquad (1.2.5)$$

prescribed on an open subset $\mathscr{O} \subset \{0\} \times \mathbb{R}^n \subset \mathbb{R} \times \mathbb{R}^n$ there exists a unique solution g^{AB} defined on an open neighborhood $\mathscr{U} \subset \mathbb{R} \times \mathbb{R}^n$ of $\mathscr{O}$ of (1.2.4). The set $\mathscr{U}$ can be chosen so that g^{AB} defines a Lorentzian metric, with $(\mathscr{U}, g)$ — globally hyperbolic with Cauchy surface $\mathscr{O}$.

Remark 1.2.2 The results in [98–100, 152] and references therein allow one to reduce the differentiability threshold above.

It remains to find out how to ensure the conditions (1.1.17). The key observation of Yvonne Choquet-Bruhat is that (1.2.4) and the Bianchi identities imply a wave equation for λ^A's. In order to see that, recall that it follows from the Bianchi identities that the Ricci tensor of the metric g necessarily satisfies a divergence identity:

$$\nabla_A\Big(R^{AB} - \frac{R}{2}g^{AB}\Big) = 0 .$$

Assuming that (1.2.4) holds, (1.2.1) implies then

$$\begin{aligned} 0 &= -\nabla_A\Big(\nabla^A\lambda^B + \nabla^B\lambda^A - \nabla_C\lambda^C g^{AB}\Big) \\ &= -\Big(\Box\lambda^B + \nabla_A\nabla^B\lambda^A - \nabla^B\nabla_C\lambda^C\Big) \\ &= -\Big(\Box\lambda^B + R^B{}_A\lambda^A\Big)\ . \end{aligned} \tag{1.2.6}$$

This shows that λ^A necessarily satisfies the second order hyperbolic system of equations

$$\Box\lambda^B + R^B{}_A\lambda^A = 0\ . \tag{1.2.7}$$

Now, it is a standard fact in the theory of hyperbolic equations that we will have

$$\lambda^A \equiv 0$$

on the domain of dependence $\mathscr{D}(\mathscr{O})$ provided that both λ^A and its derivatives vanish at $\mathscr{O}$.

REMARK 1.2.3 Actually the vanishing of $\lambda := (\lambda^A)$ as above is a completely standard result only if the metric is $C^{1,1}$; this is proved by a simpler version of the argument that we are about to present. But the result remains true under the weaker conditions of Theorem 1.2.1, which can be seen as follows. Consider initial data as in (1.2.5), with some $k \in \mathbb{R}$ satisfying $k > n/2$. Then the derivatives of the metric are in L^∞,

$$|\partial g| \le C\ ,$$

for some constant C which, in the calculation below, might change from line to line. Let $\mathscr{S}_t$ be a foliation by spacelike hypersurfaces of a conditionally compact domain of dependence $\mathscr{D}(\mathscr{S}_0)$, where $\mathscr{S}_0$ is a subset of the initial data surface $\mathscr{S}$. When λ vanishes at $\mathscr{S}_0$, a standard energy calculation for (1.2.7) gives the inequality

$$\begin{aligned} \|\lambda\|^2_{H^1(\mathscr{S}_t)} &\le C\int_0^t \|\big((1+|\mathrm{Ric}|)|\lambda| + (1+|\partial g|)|\partial\lambda|\big)|\partial\lambda|\|_{L^1(\mathscr{S}_s)}ds \\ &\le C\int_0^t \big(\|(1+|\mathrm{Ric}|)\lambda\|_{L^2(\mathscr{S}_s)}\|\partial\lambda\|_{L^2(\mathscr{S}_s)} + \|\lambda\|^2_{H^1(\mathscr{S}_t)}\big)ds \\ &\le C\int_0^t \big(\|(1+|\mathrm{Ric}|)\lambda\|_{L^2(\mathscr{S}_s)}\|\lambda\|_{H^1(\mathscr{S}_s)} + \|\lambda\|^2_{H^1(\mathscr{S}_s)}\big)ds\ . \end{aligned} \tag{1.2.8}$$

We want to use this inequality to show that λ vanishes everywhere; the idea is to estimate the integrand by a function of $\|\lambda\|^2_{H^1(\mathscr{S}_s)}$, the vanishing of λ follows then from the Gronwall lemma. Such an estimate is clear from (1.2.8) if $|\mathrm{Ric}|$ is in L^∞, which proves the claim for metrics in $C^{1,1}$, but is not obviously apparent for less regular metrics. Now, the construction of g in the course of the proof of Theorem 1.2.1 provides a metric such that $\partial g|_{\mathscr{S}_s} \in H^k$ and $\mathrm{Ric}|_{\mathscr{S}_s} \in H^{k-1}$. By Sobolev embedding for $n > 2$ we have [10]

$$\|\lambda\|_{L^p(\mathscr{S}_s)} \le C\|\lambda\|_{H^1(\mathscr{S}_s)}\ ,$$

where $p = 2n/(n-2)$. We can thus use Hölder's inequality to obtain

$$\||\mathrm{Ric}|\lambda\|_{L^2(\mathscr{S}_s)} \le \|\mathrm{Ric}\|_{L^n(\mathscr{S}_s)}\|\lambda\|_{L^p(\mathscr{S}_s)} \le C\|\mathrm{Ric}\|_{L^n(\mathscr{S}_s)}\|\lambda\|_{H^1(\mathscr{S}_s)}\ .$$

Equation (1.2.8) gives thus

$$\|\lambda\|^2_{H^1(\mathscr{S}_t)} \quad \leq \quad C\int_0^t \big(1+\|\mathrm{Ric}\|_{L^n(\mathscr{S}_s)}\big)\|\lambda\|^2_{H^1(\mathscr{S}_s)}ds\ ,$$

which is the desired inequality provided that $\|\mathrm{Ric}\|_{L^n(\mathscr{S}_s)}$ is finite. But, again by Sobolev,

$$\|\mathrm{Ric}\|_{L^p(\mathscr{S}_s)} \leq C\|\mathrm{Ric}\|_{H^{k-1}(\mathscr{S}_s)} \ \text{ provided that } \ \frac{1}{p}\geq\frac{1}{2}-\frac{k-1}{n}\ ,$$

and we see that $\mathrm{Ric}\in L^n(\mathscr{S}_s)$ will hold for $k>n/2$, as assumed in Theorem 1.2.1.

Remark 1.2.4 There exists a simple generalization of the wave coordinates condition $\Box_g x^\mu = 0$ to

$$\Box_g y^A = \mathring{\lambda}^A(y^B,x^\mu,g_{\alpha\beta})\ . \tag{1.2.9}$$

In lieu of solving the equation $\hat{E}^{AB}=0$ one solves

$$\hat{E}^{AB} = \nabla^A\mathring{\lambda}^B+\nabla^B\mathring{\lambda}^A\ . \tag{1.2.10}$$

There exists a variation of Theorem 1.2.1 that applies to this equation as well. Equation (1.2.3) can then be rewritten as

$$R^{AB}=\frac{1}{2}(\underbrace{\hat{E}^{AB}-\nabla^A\mathring{\lambda}^B-\nabla^B\mathring{\lambda}^A}_{=0})-\nabla^A(\lambda^B-\mathring{\lambda}^B)-\nabla^B(\lambda^A-\mathring{\lambda}^A)+\frac{2\Lambda}{n-1}g^{AB}\ . \tag{1.2.11}$$

This allows one to repeat the calculation (1.2.6), with λ^A there replaced by $\lambda^A-\mathring{\lambda}^A$.

There remains the easy task to adapt the calculations that follow, done in the case $\mathring{\lambda}^A=0$, to the modified condition (1.2.9), leading to initial data satisfying the right conditions.

Remark 1.2.5 We can further generalize to include matter fields. Consider, for example, a set of fields ψ^I, $i=1,\ldots,\mathring{N}$, for some $\mathring{N}\in\mathbb{N}$, satisfying a system of equations of the form

$$\Box_g\psi^I = F^I(\psi^J,\partial\psi^J,g,\partial g)\ . \tag{1.2.12}$$

We assume that there exists an associated energy-momentum tensor

$$T_{\mu\nu}(\psi^J,\partial\psi^J,g,\partial g)$$

which is identically divergence-free when (1.2.12) hold:

$$\nabla_\mu T^{\mu\nu}=0\ .$$

Allowing (1.2.9), instead of solving the equation $\hat{E}^{AB}=0$ one solves

$$\hat{E}^{AB}=\nabla^A\mathring{\lambda}^B+\nabla^B\mathring{\lambda}^A+16\pi\frac{G}{c^4}\Big(T^{AB}-\frac{1}{n-1}g^{CD}T_{CD}g^{AB}\Big)\ . \tag{1.2.13}$$

Theorem 1.2.1 applies to this equation as well. Equation (1.2.3) can then be rewritten as

$$\begin{aligned} R^{AB} \quad = \quad & \frac{1}{2}\underbrace{\Big(\hat{E}^{AB}-\nabla^A\mathring{\lambda}^B-\nabla^B\mathring{\lambda}^A-16\pi\frac{G}{c^4}\Big(T^{AB}-\frac{1}{n-1}g^{CD}T_{CD}g^{AB}\Big)\Big)}_{=0} \\ & -\nabla^A(\lambda^B-\mathring{\lambda}^B)-\nabla^B(\lambda^A-\mathring{\lambda}^A) \\ & +\frac{2\Lambda}{n-1}g^{AB}+8\pi\frac{G}{c^4}\left(T^{AB}-\frac{1}{n-1}g^{CD}T_{CD}g^{AB}\right)\ . \end{aligned} \tag{1.2.14}$$

Because T^{AB} has identically vanishing divergence by hypothesis, one can again repeat the calculation (1.2.6), with λ^A there replaced by $\lambda^A - \mathring{\lambda}^A$. As before, the right initial data will lead to a solution with $\lambda^A = \mathring{\lambda}^A$, and hence to the desired solution of the Einstein equations with sources.

We return to the vanishing of λ^A and its derivatives on $\mathscr{S}$. It is convenient to assume that y^0 is the coordinate along the $\mathbb{R}$ factor of $\mathbb{R}\times\mathbb{R}^n$, so that set $\mathscr{O}$ carrying the initial data is a subset of $\{y^0 = 0\}$; this can always be done. We have

$$\begin{aligned}\Box y^A &= \frac{1}{\sqrt{|\det g|}}\partial_B\Big(\sqrt{|\det g|}g^{BC}\partial_C y^A\Big)\\ &= \frac{1}{\sqrt{|\det g|}}\partial_B\Big(\sqrt{|\det g|}g^{BA}\Big)\,.\end{aligned}$$

So $\Box y^A$ will vanish at the initial data surface if and only if certain time derivatives of the metric are prescribed in terms of the space ones:

$$\partial_0\Big(\sqrt{|\det g|}g^{0A}\Big) = -\partial_i\Big(\sqrt{|\det g|}g^{iA}\Big)\,. \tag{1.2.15}$$

This implies that the initial data (1.2.5) for the equation (1.2.4) cannot be chosen arbitrarily if we want both (1.2.4) and the Einstein equation to be simultaneously satisfied.

It should be emphasized that there is considerable freedom in choosing the wave coordinates, which is reflected in the freedom to adjust the initial values of g^{0A}'s. A popular choice is to require that on the initial hypersurface $\{y^0 = 0\}$ we have

$$g^{00} = -1\,,\qquad g^{0i} = 0\,, \tag{1.2.16}$$

and this choice simplifies the algebra considerably. (We show that (1.2.16) can always be imposed in Proposition 1.4.1 below.) Equation (1.2.15) determines then the time derivatives $\partial_0 g^{0A}|_{\{y^0=0\}}$ needed in Theorem 1.2.1, once $g_{ij}|_{\{y^0=0\}}$ and $\partial_0 g_{ij}|_{\{y^0=0\}}$ are given. So, from this point of view, the essential initial data for the evolution problem become the space metric

$$g := g_{ij}dy^i dy^j\,,$$

together with its time derivatives.

It turns out that further constraints arise from the requirement of the vanishing of the derivatives of λ. Supposing that (1.2.15) holds on $\{y^0 = 0\}$ — equivalently, supposing that λ vanishes on $\{y^0 = 0\}$, we then have

$$\partial_i\lambda^A = 0$$

on $\{y^0 = 0\}$, where the index i is used to denote tangential derivatives. In order that all derivatives vanish initially it remains to ensure that some transverse derivative does. A transverse direction is provided by the field N of unit timelike normals to $\{y^0 = 0\}$ and, as we are about to show, the vanishing of $\nabla_N\lambda$ can be expressed as

$$\Big(G_{\mu\nu} + \Lambda g_{\mu\nu}\Big)N^\mu = 0\,. \tag{1.2.17}$$

For this, it is most convenient to use an ON frame e_μ, with $e_0 = N$. It follows from the equation $E_{AB} = 0$ and (1.2.1) that

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = -\Big(\nabla_\mu \lambda_\nu + \nabla_\nu \lambda_\mu - \nabla^\alpha \lambda_\alpha g_{\mu\nu}\Big) \;,$$

which gives

$$\begin{aligned} -\Big(G_{\mu\nu} + \Lambda g_{\mu\nu}\Big) N^\mu N^\nu &= 2\nabla_0\lambda_0 - \nabla^\alpha \lambda_\alpha \underbrace{g_{00}}_{=-1} \\ &= 2\nabla_0\lambda_0 + (-\nabla_0\lambda_0 + \underbrace{\nabla_i\lambda_i}_{=0}) \\ &= \nabla_0\lambda_0 \;, \end{aligned} \tag{1.2.18}$$

which shows that the vanishing of $\nabla_0\lambda_0$ is equivalent to the vanishing of the $\mu = 0$ component in (1.2.17). Finally

$$\begin{aligned} -\Big(G_{i0} + \Lambda g_{i0}\Big) &= \underbrace{\nabla_i\lambda_0}_{=0} + \nabla_0\lambda_i - \nabla^\alpha\lambda_\alpha \underbrace{g_{i0}}_{=0} \\ &= \nabla_0\lambda_i \;, \end{aligned} \tag{1.2.19}$$

as desired.

Equations (1.2.17) are called the *general relativistic constraint equations.* We will shortly see that (1.2.15) has quite a different character from (1.2.17); the former will be referred to as a *gauge equation.*

Summarizing, we have proved:

THEOREM 1.2.6 *Under the hypotheses of Theorem 1.2.1, suppose that the initial data (1.2.5) satisfy (1.2.15), (1.2.16) as well as the constraint equations (1.2.17). Then the metric given by Theorem 1.2.1 on the globally hyperbolic set* $\mathscr{U}$ *satisfies the vacuum Einstein equations.*

In conclusion, in the wave gauge $\lambda^A = 0$ the Cauchy data for the vacuum Einstein equations consist of

1. An open subset $\mathscr{O}$ of $\mathbb{R}^n$,
2. together with matrix-valued functions g^{AB}, $\partial_0 g^{AB}$ prescribed there, so that g^{AB} is symmetric with signature $(-,+,\cdots,+)$ at each point.
3. The constraint equations (1.2.17) hold, and
4. the algebraic gauge equation (1.2.15) holds.

So far we have been using the notation y^A for the wave coordinates. Let us assume that those coordinates are used, and let us revert to our standard notation, x^μ, for the local coordinates. In this notation, (1.1.20) can be rewritten as

$$E^{\alpha\beta} = \Box_g g^{\alpha\beta} - 2g^{\gamma\delta} g^{\epsilon\phi} \Gamma^\alpha_{\gamma\epsilon} \Gamma^\beta_{\delta\phi} - \frac{4\Lambda}{n-1} g^{\alpha\beta} \tag{1.2.20}$$

(recall that we want this to be zero in vacuum). Set

$$\varphi := \sqrt{|\det g_{\alpha\beta}|}\,, \qquad \mathfrak{g}^{\alpha\beta} := \varphi g^{\alpha\beta}\ . \tag{1.2.21}$$

In terms of $\mathfrak{g}$, the wave conditions take the particularly simple form

$$\partial_\alpha \mathfrak{g}^{\alpha\beta} = 0\ . \tag{1.2.22}$$

It is therefore convenient to rewrite Einstein equations as a system of wave equations for $\mathfrak{g}^{\alpha\beta}$. In order to do that, we calculate as follows:

$$\begin{aligned}
\partial_\mu\varphi &= \partial_\mu\left(\sqrt{|\det g_{\alpha\beta}|}\right) = \frac{1}{2}\sqrt{|\det g_{\alpha\beta}|}g^{\alpha\beta}\partial_\mu g_{\alpha\beta} = -\frac{1}{2}\sqrt{|\det g_{\alpha\beta}|}g_{\alpha\beta}\partial_\mu g^{\alpha\beta}\\
&= -\frac{1}{2}\varphi g_{\alpha\beta}\partial_\mu g^{\alpha\beta}\ ,\\
\Box_g\varphi &= \nabla^\mu\partial_\mu\varphi = -\frac{1}{2}\nabla^\mu\left(\varphi g_{\alpha\beta}\partial_\mu g^{\alpha\beta}\right)\\
&= -\frac{1}{2}\Big(\nabla^\mu\varphi\underbrace{g_{\alpha\beta}\partial_\mu g^{\alpha\beta}}_{-2\partial_\mu\varphi/\varphi} + \varphi g^{\mu\nu}\partial_\nu g_{\alpha\beta}\partial_\mu g^{\alpha\beta} + \varphi g_{\alpha\beta}\underbrace{\Box_g g^{\alpha\beta}}_{=E^{\alpha\beta}+\ldots}\Big)\\
&= \varphi^{-1}\nabla^\mu\varphi\partial_\mu\varphi - \frac{\varphi}{2}\Big(g^{\mu\nu}\partial_\nu g_{\alpha\beta}\partial_\mu g^{\alpha\beta} + g_{\alpha\beta}\big(E^{\alpha\beta} + 2g^{\gamma\delta}g^{\epsilon\phi}\Gamma^\alpha_{\gamma\epsilon}\Gamma^\beta_{\delta\phi} + \frac{4\Lambda}{n-1}g^{\alpha\beta}\big)\Big)\ ,\\
\Box_g\mathfrak{g}^{\alpha\beta} &= \varphi\Box_g g^{\alpha\beta} + 2\nabla^\mu\varphi\partial_\mu g^{\alpha\beta} + \Box_g\varphi g^{\alpha\beta}\ .
\end{aligned}$$

Thus, in harmonic coordinates,

$$\begin{aligned}
\Box_g\mathfrak{g}^{\alpha\beta} &= \varphi\big(E^{\alpha\beta} + 2g^{\gamma\delta}g^{\epsilon\phi}\Gamma^\alpha_{\gamma\epsilon}\Gamma^\beta_{\delta\phi} + \frac{4\Lambda}{n-1}g^{\alpha\beta}\big) + 2\nabla^\mu\varphi\partial_\mu g^{\alpha\beta}\\
&\quad + \left[\varphi^{-1}\nabla^\mu\varphi\partial_\mu\varphi - \frac{\varphi}{2}\Big(g^{\mu\nu}\partial_\nu g_{\rho\sigma}\partial_\mu g^{\rho\sigma} + g_{\rho\sigma}\big(E^{\rho\sigma} + 2g^{\gamma\delta}g^{\epsilon\phi}\Gamma^\rho_{\gamma\epsilon}\Gamma^\sigma_{\delta\phi} + \frac{4\Lambda}{n-1}g^{\rho\sigma}\big)\Big)\right]g^{\alpha\beta}\ ;
\end{aligned} \tag{1.2.23}$$

also note that the Λ terms can be grouped together to $-2\Lambda\mathfrak{g}^{\alpha\beta}$.

Next, it might be convenient instead to write directly equations for $g_{\mu\nu}$ rather than $g^{\mu\nu}$, or $\mathfrak{g}^{\mu\nu}$. For this, we use again $g_{\alpha\beta}g^{\beta\gamma} = \delta^\gamma_\alpha$ to obtain

$$\begin{aligned}
\partial_\sigma g_{\alpha\beta} &= -g_{\alpha\gamma}g_{\beta\delta}\partial_\sigma g^{\gamma\delta}\ ,\\
g^{\rho\sigma}\partial_\rho\partial_\sigma g_{\alpha\beta} &= -g^{\rho\sigma}\Big(\partial_\rho g_{\alpha\gamma}g_{\beta\delta}\partial_\sigma g^{\gamma\delta} + g_{\alpha\gamma}\partial_\rho g_{\beta\delta}\partial_\sigma g^{\gamma\delta}\\
&\quad + g_{\alpha\gamma}g_{\beta\delta}\partial_\rho\partial_\sigma g^{\gamma\delta}\Big)\ ,\\
&= -g^{\rho\sigma}\Big(\partial_\rho g_{\alpha\gamma}g_{\beta\delta}\partial_\sigma g^{\gamma\delta} + g_{\alpha\gamma}\partial_\rho g_{\beta\delta}\partial_\sigma g^{\gamma\delta}\Big)\\
&\quad - g_{\alpha\gamma}g_{\beta\delta}\underbrace{g^{\rho\sigma}\partial_\rho\partial_\sigma g^{\gamma\delta}}_{\Box_g g^{\gamma\delta}+\Gamma^\lambda_{\rho\sigma}\partial_\lambda g^{\gamma\delta}}\ ,\\
\Box_g g_{\alpha\beta} &= g^{\rho\sigma}\partial_\rho\partial_\sigma g_{\alpha\beta} - \Gamma^\sigma_{\alpha\beta}\partial_\sigma g_{\alpha\beta}\ .
\end{aligned}$$

One can use now the formula (1.2.20) expressing $\Box_g g^{\gamma\delta}$ in terms of $E^{\alpha\beta}$ to obtain an expression for $R_{\alpha\beta}$. In particular one finds

$$R_{\alpha\beta} = -\frac{1}{2}\Box_g g_{\alpha\beta} - g_{\alpha\mu}\nabla_\beta\lambda^\mu - g_{\beta\mu}\nabla_\alpha\lambda^\mu + \ldots\ , \tag{1.2.24}$$

where "…" stands for terms which do not involve second derivatives of the metric.

1.3 The geometry of spacelike submanifolds

Let $\mathscr{S}$ be a hypersurface in a Lorentzian or Riemannian manifold $(\mathscr{M}, g)$, we want to analyze the geometry of such hypersurfaces. Set

$$h := g|_{T\mathscr{S}} \; . \tag{1.3.1}$$

More precisely,

$$\forall \; X, Y \in T\mathscr{S} \qquad h(X,Y) := g(X,Y) \; .$$

The tensor field h is called *the first fundamental form of* $\mathscr{S}$; when non-degenerate, it is also called *the metric induced by* g *on* h. If $\mathscr{S}$ is considered as an abstract manifold with embedding $i : \mathscr{S} \to \mathscr{M}$, then h is simply the pull-back i^*g.

A hypersurface $\mathscr{S}$ will be said to be *spacelike* at $p \in \mathscr{S}$ if h is Riemannian at p, *timelike* at p if h is Lorentzian at p, and finally *null* or *isotropic* or *lightlike* at p if h is degenerate at p. $\mathscr{S}$ will be called *spacelike* if it is spacelike at all $p \in \mathscr{S}$, *etc.* An example of null hypersurface is given by $\dot{J}(p) \setminus \{p\}$ for any $p \in \mathscr{M}$, at least near p where $\dot{J}(p) \setminus \{p\}$ is differentiable.

When g is Riemannian, then h is always a Riemannian metric on $\mathscr{S}$, and then $T\mathscr{S}$ is in direct sum with $(T\mathscr{S})^\perp$. Whatever the signature of g, in this section we will always assume that this is the case:

$$T\mathscr{S} \cap (T\mathscr{S})^\perp = \{0\} \quad \Longrightarrow \quad T\mathscr{M} = T\mathscr{S} \oplus (T\mathscr{S})^\perp \; . \tag{1.3.2}$$

Recall that (1.3.2) fails precisely at those points $p \in \mathscr{S}$ at which h is degenerate. Hence, in this section we consider hypersurfaces which are either timelike throughout, or spacelike throughout. Depending upon the character of $\mathscr{S}$ we will then have

$$\epsilon := g(N,N) = \pm 1 \; , \tag{1.3.3}$$

where N is the field of unit normals to $\mathscr{S}$.

For $p \in T\mathscr{S}$ let $P : T_p\mathscr{M} \to T_p\mathscr{M}$ be defined as

$$T_p\mathscr{M} \ni X \to P(X) = X - \epsilon g(X,N)N \; . \tag{1.3.4}$$

We note the following properties of P:

- P annihilates N:

$$P(N) = P(N - \epsilon g(N,N)N) = P(N) - \epsilon^2 P(N) = 0 \; .$$

- P is a projection operator:

$$\begin{aligned} P(P(X)) &= P(X - \epsilon g(X,N)N) \\ &= P(X) - \epsilon g(X,N)P(N) = P(X) \; . \end{aligned}$$

- P restricted to $N^\perp$ is the identity:

$$g(X,N) = 0 \quad \Longrightarrow \quad P(X) = X \; .$$

- P is symmetric:

$$g(P(X),Y) = g(X,Y) - \epsilon g(X,N)g(Y,N) = g(X,P(Y)) \; .$$

The *Weingarten map* $B : T\mathscr{S} \to T\mathscr{S}$ is defined by the equation

$$T\mathscr{S} \ni X \quad \to \quad B(X) := P(\nabla_X N) \in T\mathscr{S} \subset T\mathscr{M} \; . \tag{1.3.5}$$

Here, and in other formulae involving differentiation, one should in principle choose an extension of N off $\mathscr{S}$; however, (1.3.5) involves only derivatives in directions tangent to $\mathscr{S}$, so that the result will not depend upon that extension.

In fact, the projector P is not needed in (1.3.5):

$$P(\nabla_X N) = \nabla_X N \; .$$

This follows from the calculation

$$0 = X(\underbrace{g(N,N)}_{0}) = 2g(\nabla_X N, N) \; ,$$

which shows that $\nabla_X N$ is orthogonal to N, hence tangent to $\mathscr{S}$.

The map B is closely related to the *second fundamental form* K of $\mathscr{S}$, also called the *extrinsic curvature tensor* in the physics literature:

$$T\mathscr{S} \ni X,Y \quad \to \quad K(X,Y) := g(P(\nabla_X N), Y) \tag{1.3.6a}$$
$$= g(B(X),Y) \; . \tag{1.3.6b}$$

It is often convenient to have at our disposal index formulae, for this purpose let us consider a local ON frame $\{e_\mu\}$ such that $e_0 = N$ along $\mathscr{S}$. We then have

$$g_{\mu\nu} = \mathrm{diag}(-1, +1, \ldots, +1)$$

in the case of a spacelike hypersurface in a Lorentzian manifold.

Using the properties of P listed above,

$$\begin{aligned} K_{ij} &:= K(e_i,e_j) = g(P(\nabla_{e_i}N), e_j) = g(P(P(\nabla_{e_i}N)), e_j) \\ &= g(P(\nabla_{e_i}N), P(e_j)) = h(P(\nabla_{e_i}N), e_j) = h(B^k{}_i e_k, e_j) \\ &= h_{kj} B^k{}_i \; , \end{aligned} \tag{1.3.7}$$

$$B^k{}_i := \varphi^k(B(e_i)) \; , \tag{1.3.8}$$

where $\{\varphi^k\}$ is a basis of $T^*\mathscr{S}$ dual to the basis $\{P(e_i)\}$ of $T\mathscr{S}$. Equivalently,

$$B^k{}_i = h^{kj} K_{ji} \; ,$$

and it is usual to write the right-hand side as $K^k{}_i$.

Let us show that K is symmetric: first,

$$\begin{aligned} K(X,Y) &= g(\nabla_X N, Y) \\ &= X(\underbrace{g(N,Y)}_{=0}) - g(N, \nabla_X Y) \; . \end{aligned} \tag{1.3.9}$$

Now, ∇ has no torsion, which implies

$$\nabla_X Y = \nabla_Y X - [X, Y] \ .$$

Further, the commutator of vector fields tangent to $\mathscr{S}$ is a vector field tangent to $\mathscr{S}$, which implies

$$\forall\ X, Y \in T\mathscr{S} \qquad g(N, [X, Y]) = 0 \ .$$

Returning to (1.3.9), it follows that

$$K(X, Y) = -g(N, \nabla_Y X - [X, Y]) = -g(N, \nabla_Y X) \ ,$$

and the equation

$$K(X, Y) = K(Y, X)$$

immediately follows from (1.3.9).

To continue, for X, Y — sections of $T\mathscr{S}$ we set

$$D_X Y := P(\nabla_X Y) \ . \tag{1.3.10}$$

First, we claim that D is a connection: Linearity with respect to addition in all variables, and with respect to multiplication of X by a function, is straightforward. It remains to check the Leibniz rule:

$$\begin{aligned} D_X(\alpha Y) &= P(\nabla_X(\alpha Y)) \\ &= P(X(\alpha)Y + \alpha \nabla_X Y) \\ &= X(\alpha)P(Y) + \alpha P(\nabla_X Y) \\ &= X(\alpha)Y + \alpha D_X Y \ . \end{aligned}$$

It follows that all the axioms of a covariant derivative on vector fields are fulfilled, as desired. It turns out that D is actually the Levi-Civita connection of the metric h. Recall that the Levi-Civita connection is determined uniquely by the requirement of vanishing torsion, and that of metric-compatibility. Both results are straightforward:

$$D_X Y - D_Y X = P(\nabla_X Y - \nabla_Y X) = P([X, Y]) = [X, Y] \ ;$$

in the last step we have again used the fact that the commutator of two vector fields tangent to $\mathscr{S}$ is a vector field tangent to $\mathscr{S}$. In order to establish metric-compatibility, we calculate for all vector fields X, Y, Z tangent to $\mathscr{S}$:

$$\begin{aligned} X(h(Y, Z)) &= X(g(Y, Z)) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ &= g(\nabla_X Y, \underbrace{P(Z)}_{=Z}) + g(\underbrace{P(Y)}_{=Y}, \nabla_X Z) \\ &= \underbrace{g(P(\nabla_X Y), Z) + g(Y, P(\nabla_X Z))}_{P \text{ is symmetric}} \\ &= g(D_X Y, Z) + g(Y, D_X Z) \\ &= h(D_X Y, Z) + h(Y, D_X Z) \ . \end{aligned}$$

Equation (1.3.10) turns out to be very convenient when trying to express the curvature of h in terms of that of g. To distinguish between both curvatures let us use the symbol ρ for the curvature tensor of h; by definition, for all vector fields tangential to $\mathscr{S}$,

$$\begin{aligned}\rho(X,Y)Z &= D_X D_Y Z - D_Y D_X Z - D_{[X,Y]}Z \\ &= P\Big(\nabla_X(P(\nabla_Y Z)) - \nabla_Y(P(\nabla_X Z)) - \nabla_{[X,Y]}Z\Big)\,.\end{aligned}$$

Now, for any vector field W (Not necessarily tangent to $\mathscr{S}$) we have

$$\begin{aligned}P\Big(\nabla_X(P(W))\Big) &= P\Big(\nabla_X(W - \epsilon g(N,W)N)\Big) \\ &= P\Big(\nabla_X W - \underbrace{\epsilon X(g(N,W))n}_{P(N)=0} - \epsilon g(N,W)\nabla_X n\Big) \\ &= P\Big(\nabla_X W\Big) - \epsilon g(N,W)P\Big(\nabla_X n\Big) \\ &= P\Big(\nabla_X W\Big) - \epsilon g(N,W)B(X)\,.\end{aligned}$$

Applying this equation to $W = \nabla_Y Z$ we obtain

$$\begin{aligned}P\Big(\nabla_X(P(\nabla_Y Z))\Big) &= P(\nabla_X\nabla_Y Z) - \epsilon g(N,\nabla_Y Z)B(X) \\ &= P(\nabla_X\nabla_Y Z) + \epsilon K(Y,Z)B(X)\,,\end{aligned}$$

and in the last step we have used (1.3.9). It now immediately follows that

$$\rho(X,Y)Z = P(R(X,Y)Z) + \epsilon\Big(K(Y,Z)B(X) - K(X,Z)B(Y)\Big)\,. \tag{1.3.11}$$

In an adapted ON frame as discussed above this reads

$$\boxed{\rho^i{}_{jk\ell} = R^i{}_{jk\ell} + \epsilon(K^i{}_k K_{j\ell} - K^i{}_\ell K_{jk})}\,. \tag{1.3.12}$$

Here $K^i{}_k$ is the tensor field K_{ij} with an index raised using the contravariant form $h^\#$ of the metric h, compare (1.3.7).

We are ready now to derive the *general relativistic scalar constraint equation*: Let ρ_{ij} denote the Ricci tensor of the metric h, we then have

$$\begin{aligned}\rho_{j\ell} &:= \rho^i{}_{ji\ell} \\ &= \underbrace{R^i{}_{ji\ell}}_{=R^\mu{}_{j\mu\ell} - R^0{}_{j0\ell}} + \epsilon(K^i{}_i K_{j\ell} - K^i{}_\ell K_{ji}) \\ &= R_{j\ell} - R^0{}_{j0\ell} + \epsilon(\mathrm{tr}_h K K_{j\ell} - K^i{}_\ell K_{ji})\,.\end{aligned}$$

Defining $R(h)$ to be the scalar curvature of h, it follows that

$$\begin{aligned}R(h) &= \rho^j{}_j \\ &= \underbrace{R^j j}_{=R^\mu{}_\mu - R^0{}_0} - \underbrace{R^{0j}{}_{0j}}_{=R^{0\mu}{}_{0\mu}} + \epsilon(\mathrm{tr}_h K K^j{}_j - K^{ij}K_{ji}) \\ &= R(g) - 2\underbrace{R^0{}_0}_{=\epsilon R_{00}} + \epsilon\Big((\mathrm{tr}_h K)^2 - |K|_h^2\Big) \\ &= -16\pi\epsilon T_{00} + 2\Lambda + \epsilon\Big((\mathrm{tr}_h K)^2 - |K|_h^2\Big)\,,\end{aligned}$$

and we have used the Einstein equation,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi\frac{G}{c^4}T_{\mu\nu} \,, \tag{1.3.13}$$

with $G = c = 1$. Assuming that $\epsilon = -1$ we obtain the desired scalar constraint:

$$\boxed{R(h) = 16\pi T_{\mu\nu}N^\mu N^\nu + 2\Lambda + |K|_h^2 - (\mathrm{tr}_h K)^2} \,. \tag{1.3.14}$$

(We emphasise that this equation is valid whatever the dimension of $\mathscr{S}$.) In particular in vacuum, with $\Lambda = 0$, one obtains

$$R(h) = |K|_h^2 - (\mathrm{tr}_h K)^2 \,. \tag{1.3.15}$$

The *vector constraint equation* carries the remaining information contained in the equation $G_{\mu\nu}N^\mu = 0$. In order to understand that equation let Y be tangent to $\mathscr{S}$, we then have

$$\begin{aligned} G_{\mu\nu}N^\mu Y^\nu &= \Big(R_{\mu\nu} - \frac{1}{2}R(g)g_{\mu\nu}\Big)N^\mu Y^\nu \\ &= \mathrm{Ric}(N,Y) - \frac{1}{2}R(g)g(N,Y) \\ &= \mathrm{Ric}(N,Y) \,. \end{aligned} \tag{1.3.16}$$

We will relate this to some derivatives of K. By definition we have

$$(D_Z K)(X,Y) = Z(K(X,Y)) - K(D_Z X, Y) - K(X, D_Z Y) \,.$$

Now,

$$Z(K(X,Y)) \;=\; Z(g(\nabla_X N, Y)) = g(\nabla_Z\nabla_X N, Y) + g(\nabla_X N, \nabla_Z Y) \,.$$

Since $\nabla_X N$ is tangential, and P is symmetric, the last term can be rewritten as

$$\begin{aligned} g(\nabla_X N, \nabla_Z Y) &= g(P(\nabla_X N), \nabla_Z Y) = g(\nabla_X N, P(\nabla_Z Y)) \\ &= K(X, P(\nabla_Z Y)) = K(X, D_Z Y) \,. \end{aligned}$$

It follows that

$$\begin{aligned} &(D_Z K)(X,Y) - (D_X K)(Z,Y) \\ &\quad= g(\nabla_Z\nabla_X N, Y) - g(\nabla_X\nabla_Z n, Y) + K(X, D_Z Y) - K(Z, D_X Y) \\ &\qquad -K(D_Z X, Y) - K(X, D_Z Y) + K(D_X Z, Y) + K(Z, D_X Y) \\ &\quad= g(R(Z,X)N,Y) + \underbrace{g(\nabla_{[Z,X]}n, Y)}_{K([Z,X],Y)} - K(\underbrace{D_Z X - D_X Z}_{[Z,X]}, Y) \\ &\quad= g(R(Z,X)N,Y) \,. \end{aligned}$$

Thus,

$$(D_Z K)(X,Y) - (D_X K)(Z,Y) = g(R(Z,X)N,Y) \,. \tag{1.3.17}$$

In a frame in which the e_i's are tangent to the hypersurface $\mathscr{S}$, this can be rewritten as

$$D_k K_{ij} - D_i K_{kj} = R_{j\mu ki} N^\mu \ . \tag{1.3.18}$$

A contraction over i and j gives then

$$h^{ij}(D_k K_{ij} - D_i K_{kj}) = h^{ij} R_{j0ki} + \epsilon \underbrace{R_{00k0}}_{0} = g^{\mu\nu} R_{\mu 0 k\nu} = -R_{k0} \ .$$

Using the Einstein equation (1.3.13) together with (1.3.16) we obtain the *vector constraint equation*:

$$\boxed{D_j K^j{}_k - D_k K^j{}_j = 8\pi T_{\mu\nu} N^\mu h^\nu{}_k \ .} \tag{1.3.19}$$

1.4 Cauchy data

Let us return to the discussion of the end of Section 1.1. We shall adopt a slightly general point of view than that presented there, where we assumed that the initial data were given on an open subset $\mathscr{O}$ of the zero-level set of the function y^0. A correct geometric picture here is to start with an n-dimensional hypersurface $\mathscr{S}$, and prescribe initial data there; the case where $\mathscr{S}$ is $\mathscr{O}$ is thus a special case of this construction. At this stage there are two attitudes one may wish to adopt: the first is that $\mathscr{S}$ is a subset of the space-time $\mathscr{M}$ — this is essentially what we assumed in Section 1.3. Another way of looking at this is to consider $\mathscr{S}$ as a hypersurface of its own, equipped with an embedding

$$i : \mathscr{S} \to \mathscr{M} \ .$$

The most convenient approach is to go back and forth between those points of view, and this is the strategy that we will follow.

As made clear by the results in Section 1.3, the metric h is uniquely defined by the space-time metric g once that $\mathscr{S} \subset \mathscr{M}$ (or $i(\mathscr{S}) \subset \mathscr{M}$) has been prescribed; the same applies to the extrinsic curvature tensor K. A *vacuum initial data set* $(\mathscr{S}, h, K)$ is a triple where $\mathscr{S}$ is an n–dimensional manifold, h is a Riemannian metric on $\mathscr{S}$, and K is a symmetric two-covariant tensor field on $\mathscr{S}$. Further (h, K) are supposed to satisfy the vacuum constraint equations (1.3.15) and (1.3.19), perhaps (but not necessarily so) with a non-zero cosmological constant Λ.

Let us show that specifying K is equivalent to prescribing the time-derivatives of the space-part g_{ij} of the resulting space-time metric g. Suppose, indeed, that a space-time (M, g) has been constructed (not necessarily vacuum) such that K is the extrinsic curvature tensor of $\mathscr{S}$ in $(\mathscr{M}, g)$. Consider any domain of coordinates $\mathscr{O} \subset \mathscr{S}$ and construct coordinates y^μ in some $\mathscr{M}$–neighborhood of $\mathscr{U}$ such that $\mathscr{S} \cap \mathscr{U} = \mathscr{O}$; those coordinates could be wave-coordinates, as described at the end of Section 1.1, but this is not necessary at this stage. Since y^0 is constant on $\mathscr{S}$ the one-form dy^0 annihilates $T\mathscr{S}$, so does the one form $g(N, \cdot)$, and since $\mathscr{S}$ has codimension one it follows that dy^0 must be proportional to $g(N, \cdot)$:

$$N_A dy^A = N_0 dy^0$$

on $\mathscr{O}$. The normalization $-1 = g(n,n) = g^{\mu\nu}n_\mu n_\nu = g^{00}(n_0)^2$ gives

$$n_\alpha dy^\alpha = \frac{1}{\sqrt{|g^{00}|}} dy^0 \ .$$

Next,

$$\begin{aligned} K_{ij} &:= g(\nabla_i N, \partial_j) = \nabla_i N_j \\ &= \partial_i N_j - \Gamma^\mu{}_{ji} N_\mu \\ &= -\Gamma^0_{ji} N_0 \\ &= -\frac{1}{2} g^{0\sigma}\Big(\partial_j g_{\sigma i} + \partial_i g_{\sigma j} - \partial_\sigma g_{ij}\Big) N_0 \ . \end{aligned} \qquad (1.4.1)$$

This shows that the knowledge of $g_{\mu\nu}$ *and* $\partial_0 g_{ij}$ at $y^0 = 0$ allows one to calculate K_{ij}. Reciprocally, (1.4.1) can be rewritten as

$$\partial_0 g_{ij} = \frac{2}{g^{00} N_0} K_{ij} + \text{ terms determined by the } g_{\mu\nu}\text{'s and their space–derivatives} \ ,$$

so that the knowledge of the $g_{\mu\nu}$'s and of the K_{ij}'s at $y^0 = 0$ allows one to calculate $\partial_0 g_{ij}$. Thus, K_{ij} is the geometric counterpart of the $\partial_0 g_{ij}$'s.

It is sometimes said that the g_{0A}'s have a *gauge character*. By this it is usually meant that the objects under consideration do not have any intrinsic meaning, and their values can be changed using the action of some family of transformations, relevant to the problem at hand, without changing the geometric, or physical, information carried by those objects. In our case the relevant transformations are the coordinate ones, and things are made precise by the following proposition:

PROPOSITION 1.4.1 *Let g_{AB}, $\tilde g_{AB}$ be two metrics such that*

$$g_{ij}|_{\{y^0=0\}} = \tilde g_{ij}|_{\{y^0=0\}} \ , \qquad K_{ij}|_{\{y^0=0\}} = \tilde K_{ij}|_{\{y^0=0\}} \ . \qquad (1.4.2)$$

Then there exists a coordinate transformation ϕ defined in a neighborhood of $\{y^0 = 0\}$ which preserves (1.4.2) such that

$$g_{0A}|_{\{y^0=0\}} = (\phi^*\tilde g)_{0A}|_{\{y^0=0\}} \ . \qquad (1.4.3)$$

Furthermore, for any metric g there exist local coordinate systems $\{\bar y^\mu\}$ such that $\{y^0 = 0\} = \{\bar y^0 = 0\}$ and, if we write $g = \bar g_{AB} d\bar y^A d\bar y^B$ etc. *in the barred coordinate system, then*

$$\begin{aligned} g_{ij}|_{\{y^0=0\}} = \bar g_{ij}|_{\{\bar y^0=0\}} \ , \quad K_{ij}|_{\{y^0=0\}} = \bar K_{ij}|_{\{\bar y^0=0\}} \ , \\ \bar g_{00}|_{\{y^0=0\}} = -1 \ , \quad \bar g_{0i}|_{\{y^0=0\}} = 0 \ . \end{aligned} \qquad (1.4.4)$$

REMARK 1.4.2 We can actually always achieve $\bar g_{00} = -1$, $\bar g_{0i} = 0$ in a whole neighborhood of $\mathscr{S}$: this is done by shooting geodesics normally to $\mathscr{S}$, choosing y^0 to be the affine parameter along those geodesics, and by transporting the coordinates y^i from $\mathscr{S}$ by requiring them to be constant along the normal geodesics. The coordinate system will break down wherever the normal geodesics start intersecting, but the implicit function theorem guarantees that there will exist a neighborhood of $\mathscr{S}$ on which this does not happen. The resulting coordinates are called *Gauss coordinates*. While those coordinates are geometrically natural, in this coordinate system the Einstein equations do not appear to have good properties from the PDE point of view.

PROOF: It suffices to prove the second claim: for if $\bar{\phi}$ is the transformation that brings g to the form (1.4.4), and $\tilde{\phi}$ is the corresponding transformation for $\tilde{g}$, then $\phi := \tilde{\phi} \circ \bar{\phi}^{-1}$ will satisfy (1.4.3).

Let us start by calculating the change of the metric coefficients under a transformation of the form

$$y^0 = \varphi \bar{y}^0 \,, \qquad y^i = \bar{y}^i + \psi^i \bar{y}^0 \,. \tag{1.4.5}$$

If $\varphi > 0$ then clearly

$$\{y^0 = 0\} = \{\bar{y}^0 = 0\} \,.$$

Further, one has

$$\begin{aligned}
g\Big|_{\{y^0=0\}} &= \Big(g_{00}(dy^0)^2 + 2g_{0i}dy^0dy^i + g_{ij}dy^idy^j\Big)\Big|_{\{y^0=0\}} \\
&= \Big(g_{00}(\bar{y}^0d\varphi + \varphi d\bar{y}^0)^2 + 2g_{0i}(\bar{y}^0d\varphi + \varphi d\bar{y}^0)(d\bar{y}^i + \bar{y}^0d\psi^i + \psi^i d\bar{y}^0) \\
&\quad + g_{ij}(d\bar{y}^i + \bar{y}^0d\psi^i + \psi^i d\bar{y}^0)(d\bar{y}^j + \bar{y}^0d\psi^j + \psi^j d\bar{y}^0)\Big)\Big|_{\{y^0=0\}} \\
&= \Big(g_{00}(\varphi d\bar{y}^0)^2 + 2g_{0i}\varphi d\bar{y}^0(d\bar{y}^i + \psi^i d\bar{y}^0) \\
&\quad + g_{ij}(d\bar{y}^i + \psi^i d\bar{y}^0)(d\bar{y}^j + \psi^j d\bar{y}^0)\Big)\Big|_{\{y^0=0\}} \\
&= \Big((g_{00}\varphi^2 + 2g_{0i}\psi^i + g_{ij}\psi^i\psi^j)(d\bar{y}^0)^2 \\
&\quad + 2(g_{0i}\varphi + g_{ij}\psi^j)d\bar{y}^0d\bar{y}^i + g_{ij}d\bar{y}^id\bar{y}^j\Big)\Big|_{\{y^0=0\}} \\
&=: \bar{g}_{\mu\nu}d\bar{y}^\mu d\bar{y}^\nu \,.
\end{aligned}$$

We shall apply the above transformation twice: first we choose $\varphi = 1$ and

$$\psi^i = h^{ij}g_{0j} \,,$$

where h^{ij} is the matrix inverse to g_{ij}; this leads to a metric with $\bar{g}_{0i} = 0$. We then apply a second transformation of the form (1.4.5) to the new metric, now with the new $\psi^i = 0$, and with a φ chosen so that the final g_{00} equals minus one. □

1.5 Solutions global in space

In order to globalize the existence Theorem 1.2.1 *in space*, the key point is to show that two solutions differing only by the values $g_{0\alpha}|_{\{y^0=0\}}$ are (locally) isometric: so suppose that g and $\tilde{g}$ both solve the vacuum Einstein equations in a globally hyperbolic region $\mathscr{U}$, with the same Cauchy data (g, K) on $\mathscr{O} := \mathscr{U} \cap \mathscr{S}$. One can then introduce wave coordinates in a globally hyperbolic neighborhood of $\mathscr{O}$ both for g and $\tilde{g}$, satisfying (1.2.16), by solving

$$\Box_g y^\mu = 0 \,, \qquad \Box_{\tilde{g}} \tilde{y}^\mu = 0 \,, \tag{1.5.1}$$

with the same initial data for y^μ and $\tilde{y}^\mu$. Transforming both metrics to their respective wave-coordinates, one obtains two solutions of the reduced equation (1.1.20) with the same initial data.

The question then arises whether the resulting metrics will be sufficiently differentiable to apply the uniqueness part of Theorem 1.2.1. Now, the metrics obtained so far are in a space $C^1([0,T],H^s)$, where the Sobolev space H^s involves the space-derivatives of the metric. The initial data for the solutions y^μ or $\tilde{y}^\mu$ of (1.5.1) may be chosen to be in $H^{s+1}\times H^s$. However, a rough inspection of (1.5.1) shows that the resulting solutions will be only in $C^1([0,T],H^s)$, because of the low regularity of the metric. But then (1.1.8) implies that the transformed metrics will be in $C^1([0,T],H^{s-1})$, and uniqueness can only be invoked *provided that* $s-1>n/2+1$, which is one degree of differentiability more than what was required for existence. This was the state of affairs for some fifty-five years until the following simple argument of Planchon and Rodnianski [137]: To make it clear that the functions y^μ are considered to be scalars in (1.5.1), we shall write y for y^μ. Commuting derivatives with $\Box_g$ one finds, for metrics satisfying the vacuum Einstein equations,

$$\Box_g \nabla_\alpha y = \nabla_\mu \nabla^\mu \nabla_\alpha y = [\nabla_\mu \nabla^\mu, \nabla_\alpha] y = \underbrace{R^{\sigma\mu}{}_{\alpha\mu}}_{=R^\sigma{}_\alpha=0} \nabla_\sigma y = 0 \; .$$

Commuting once more one obtains an evolution equation for the field $\psi_{\alpha\beta} := \nabla_\alpha \nabla_\beta y$:

$$\Box_g \psi_{\alpha\beta} + \underbrace{\nabla_\sigma R_\beta{}^\lambda{}_\alpha{}^\sigma}_{=0} \nabla_\lambda y + 2R_\beta{}^\lambda{}_\alpha{}^\sigma \psi_{\sigma\lambda} = 0 \; ,$$

where the underbraced term vanishes, for vacuum metrics, by a contracted Bianchi identity. So the most offending term in this equation for $\psi_{\alpha\beta}$, involving three derivatives of the metric, disappears when the metric is vacuum. Standard theory of hyperbolic PDEs shows now that the functions $\nabla_\alpha\nabla_\beta y$ are in $C^1([0,T],H^{s-1})$, hence $y\in C^1([0,T],H^{s+1})$, and the transformed metrics are regular enough to invoke uniqueness without having to increase s.

Suppose, now, that an initial data set $(\mathscr{S},g,K)$ as in Theorem 1.2.1 is given. Covering $\mathscr{S}$ by coordinate neighborhoods $\mathscr{O}_p$, $p\in\mathscr{S}$, one can use Theorem 1.2.1 to construct globally hyperbolic developments $(\mathscr{U}_p,g_p)$ of $(\mathscr{U}_p,g,K)$. By the argument just given the metrics so obtained will coincide, after performing a suitable coordinate transformation, wherever simultaneously defined. This allows one to patch the $(\mathscr{U}_p,g_p)$'s together to a globally hyperbolic Lorentzian manifold, with Cauchy surface $\mathscr{S}$. Thus:

THEOREM 1.5.1 *Any vacuum initial data set $(\mathscr{S},g,K)$ of differentiability class $H^{s+1}\times H^s$, $s>n/2$, admits a globally hyperbolic development.*

The solutions are locally unique, in a sense made clear by the proof.

1.6 The Cauchy problem for the Einstein-Maxwell equations

The Einstein-Maxwell equations form a system of equations for the gravitational field and the Maxwell potential $A = A_\mu dx^\mu$. The electric and magnetic fields are encoded in an anti-symmetric tensor field

$$F = dA \quad \Longleftrightarrow \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu \ .$$

(In the last equality above we have used the fact that the Levi-Civita connection ∇ has no torsion.) The *Maxwell field* F is required to satisfy the *sourceless Maxwell equations*

$$\nabla^\mu F_{\mu\nu} = 0 \ . \tag{1.6.1}$$

There is no natural *space-time* decomposition of F into electric and magnetic parts. However, in space-time dimension four, given a (timelike) vector field T^μ normal to a family of space-like hypersurfaces $\mathscr{S}_t$ one sets

$$E_\mu dx^\mu = F_{\alpha\beta} T^\beta dx^\alpha \ , \quad B_\mu dx^\mu = \frac{1}{2}\epsilon_{\alpha\beta\gamma\delta} T^\alpha F^{\gamma\delta} dx^\beta \ . \tag{1.6.2}$$

By anti-symmetry we have

$$E_\mu T^\mu = B_\mu T^\mu = 0 \ ,$$

which shows that the pull-backs to the $\mathscr{S}_t$'s of E and B contains the whole information about E and B. Furthermore one easily checks that

$$F_{\mu\nu} = 2T_{[\mu} E_{\nu]} + \epsilon_{\mu\nu\rho\sigma} T^\rho B^\sigma \ ,$$

which shows that E and B contain the whole information about F.

The reader can check that (1.6.1) together with $dF = 0$ leads to evolution equations for the electric field E and the magnetic field B which closely resemble the Maxwell equations. Furthermore, those equations reduce to the standard Maxwell evolution equations when the space-time metric g is flat and $T^\mu \partial_\mu$ equals ∂_t.

The gravitational field is coupled to F via the Einstein equations,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu} \ , \tag{1.6.3}$$

using the following energy-momentum tensor:

$$T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\alpha} F_\nu{}^\alpha - \frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} g_{\mu\nu}\right) \ . \tag{1.6.4}$$

An interesting feature of the *Einstein-Maxwell equations* (1.6.1) and (1.6.4) is the existence of *gauge freedom*, by which one means the following: Let (g, A) be a solution of the equations, let λ be an arbitrary smooth function on space-time, and consider a new electromagnetic potential $\hat{A}$ defined as

$$\hat{A} = A + d\lambda \quad \Longleftrightarrow \quad \hat{A}_\mu = A_\mu + \partial_\mu \lambda \ . \tag{1.6.5}$$

This does not change F, since $dd\lambda = 0$, so that (1.6.1) and (1.6.3) still hold. The transformation (1.6.5) is called a *gauge transformation*. The physical interpretation is that the precise values of the potential A are irrelevant, the important field being F, so that the existence of gauge transformations does not affect the physical properties of a given solution. On the other hand it is sometimes convenient to have the electromagnetic potential field at disposal. One can supplement the above equations by various *gauge conditions* which eliminate, or reduce, the gauge freedom. One such condition which is convenient for our further purposes is the *Lorenz gauge*[1],

$$\nabla_\mu A^\mu = 0 \; . \tag{1.6.6}$$

Solutions of the Einstein-Maxwell equations can be constructed by solving a Cauchy problem, as follows: The Cauchy data consist of a gravitational initial data set

$$(\mathscr{S}, g, K)$$

together with a set of fields

$$(A_0, A_i, \partial_t A_i)$$

on $\mathscr{S}$. One seeks solutions in wave coordinates for the metric and in the Lorenz gauge for the electromagnetic potential. The evolution equation for A_μ is taken to be

$$\Box A_\mu = R_\mu{}^\nu A_\nu \; . \tag{1.6.7}$$

(This equation will be justified by the calculations in (1.6.12) below.)

In spite of apparences, the above equation does not contain second derivatives of the metric, at least in wave coordinates: the second derivatives of the metric that appear in the left-hand side through the derivatives of the Christoffel symbols cancel out exactly the ones at the right-hand side. This can be seen as follows: (1.6.7) is equivalent to $\nabla_\mu F^{\mu\nu} = 0$, which can also be rewritten as

$$\begin{aligned} 0 &= \partial_\mu \left[\sqrt{|\det g|} g^{\nu\beta} g^{\mu\alpha} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \right] \qquad (1.6.8) \\ &= g^{\nu\beta} \partial_\mu (\sqrt{|\det g|} g^{\mu\alpha} \partial_\alpha A_\beta) - g^{\nu\beta} \partial_\mu (\sqrt{|\det g|} g^{\mu\alpha} \partial_\beta A_\alpha) \\ &\quad + (\partial_\mu g^{\nu\beta}) \left[\sqrt{|\det g|} g^{\mu\alpha} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \right] \\ &= g^{\nu\beta} \sqrt{|\det g|} \Box_g A_\beta + \underbrace{g^{\nu\beta} \partial_\mu (\sqrt{|\det g|} g^{\mu\alpha}) \partial_\alpha A_\beta - g^{\nu\beta} \partial_\mu (\sqrt{|\det g|} g^{\mu\alpha}) \partial_\beta A_\alpha}_{=0 \text{ in wave coordinates}} \\ &\quad \underbrace{- g^{\nu\beta} g^{\mu\alpha} \sqrt{|\det g|} \partial_\mu \partial_\beta A_\alpha}_{(I)} + (\partial_\mu g^{\nu\beta}) \left[\sqrt{|\det g|} g^{\mu\alpha} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \right] \; , \end{aligned}$$

where $\Box_g$ is the wave operator acting on scalars, in wave coordinates $\Box_g = g^{\alpha\beta}\partial_\alpha\partial_\beta$. In Lorenz gauge we have

$$\begin{aligned} 0 &= \partial_\beta \partial_\mu (\sqrt{|\det g|} g^{\mu\alpha} A_\alpha) = \partial_\beta \Big(\underbrace{\partial_\mu (\sqrt{|\det g|} g^{\mu\alpha}) A_\alpha}_{=0 \text{ in wave coordinates}} + \sqrt{|\det g|} g^{\mu\alpha} \partial_\mu A_\alpha \Big) \\ &= \sqrt{|\det g|} g^{\mu\alpha} \partial_\beta \partial_\mu A_\alpha + \partial_\beta \left(\sqrt{|\det g|} g^{\mu\alpha} \right) \partial_\mu A_\alpha \; , \end{aligned}$$

[1]This is not a spelling error, *Ludvig Lorenz* who introduced the gauge, should not be confused with *Hendrik Lorentz* who introduced the transformations bearing his name, and who shared the 1902 Nobel prize for his explanation of the Zeeman effect.

leading to

$$
\begin{aligned}
(I) &= g^{\nu\beta}\partial_\beta\left(\sqrt{|\det g|}g^{\mu\alpha}\right)\partial_\mu A_\alpha \qquad (1.6.9)\\
&= g^{\nu\beta}\left(\frac{1}{2}\sqrt{|\det g|}g^{\lambda\tau}(\partial_\beta g_{\lambda\tau})g^{\mu\alpha}+\sqrt{|\det g|}\partial_\beta g^{\mu\alpha}\right)\partial_\mu A_\alpha \,.
\end{aligned}
$$

Finally

$$
\begin{aligned}
0 &= g^{\nu\beta}\Box_g A_\beta+\frac{1}{2}g^{\nu\beta}g^{\lambda\tau}\partial_\beta g_{\lambda\tau}g^{\mu\alpha}\partial_\mu A_\alpha+g^{\nu\beta}(\partial_\beta g^{\mu\alpha})\partial_\mu A_\alpha\\
&\quad +g^{\mu\alpha}(\partial_\mu g^{\nu\beta})(\partial_\alpha A_\beta-\partial_\beta A_\alpha)\,, \qquad (1.6.10)
\end{aligned}
$$

and we see that there are no second derivatives of the metric in this equation, as desired.

The initial data needed for (1.6.7) can be chosen to be A_μ and $\partial_t A_\mu$ on $\mathscr{S}$; the missing field $\partial_t A_0$ is calculated from the remaining ones using (1.6.6).

Given a solution of the harmonically reduced Einstein equation and of (1.6.7), we need to show that A_μ satisfies the Lorenz gauge condition (1.6.6), and that the Maxwell equations (1.6.1) hold. In order to do that, set

$$\psi := \nabla^\mu A_\mu \,, \qquad (1.6.11)$$

and we wish to show that $\psi \equiv 0$. Now,

$$
\begin{aligned}
\nabla^\mu F_{\mu\nu} &= \Box A_\mu - R_\mu{}^\nu A_\nu - \nabla_\mu\psi\\
&= -\nabla_\mu\psi \qquad (1.6.12)
\end{aligned}
$$

in view of (1.6.7). This shows that $\partial_t\psi$ will vanish on $\mathscr{S}$ if and only if we impose the *Maxwell constraint equation*

$$\nabla^\mu F_{\mu 0} = 0 \,. \qquad (1.6.13)$$

Assuming that this equation holds, we calculate

$$0 = \nabla^\mu\nabla^\nu F_{\mu\nu} = -\Box\psi \,.$$

Here we have used the fact that the left-hand-side of the last equation vanishes identically. It follows that ψ satisfies the homogeneous wave equation

$$\Box\psi = 0 \,.$$

By choice of $\partial_t A_0$ we have $\psi = 0$ on $\mathscr{S}$, while $\partial_t\psi = 0$ on $\mathscr{S}$ by the Maxwell constraint equation, hence $\psi \equiv 0$ in any globally hyperbolic development of $\mathscr{S}$. Subsequently

$$\nabla^\mu F_{\mu\nu} = 0$$

by (1.6.12). We have thus proved that the field A_μ so obtained satisfies the Maxwell equation, and is in the Lorenz gauge.

1.7 Constraint equations: the conformal method

A set (M, g, K), where (M, g) is a Riemannian manifold, and K is a symmetric tensor field on M, will be called a *vacuum initial data set* if the vacuum constraint equations (1.3.14), (1.3.19) hold:

$$D_j K^j{}_k = D_k K^j{}_j \ , \tag{1.7.1a}$$

$$R(g) = 2\Lambda + |K|_g^2 - (\mathrm{tr}_g K)^2 \ . \tag{1.7.1b}$$

Here, as before, Λ is a constant. The object of this section is to present the *conformal method* for constructing solutions of (1.7.1). This method works best when $\mathrm{tr}_g K$ is constant over M:

$$\partial_i(\mathrm{tr}_g K) = 0 \ . \tag{1.7.2}$$

(We shall see shortly that (1.7.2) leads to a decoupling of the equations (1.7.1), in a sense which will be made precise.) Hypersurfaces M in a space-times $\mathscr{M}$ satisfying (1.7.2) are known as *constant mean curvature (CMC) surfaces.* Equation (1.7.2) is sometimes viewed as a "gauge condition", in the following sense: if we require (1.7.2) to hold on all hypersurfaces M_τ within a family of hypersurfaces in the space-time, then this condition restricts the freedom of choice of the associated time function t which labels those hypersurfaces. Unfortunately there exist space-times in which no CMC hypersurfaces exist [15, 92]. Now, the conformal method is the only method known which produces *all* solutions satisfying a reasonably mild "gauge condition", it is therefore unfortunate that condition (1.7.2) is a restrictive one.

The conformal method seems to go back to Lichnerowicz [108], except that Lichnerowicz proposes a different treatment of the vector constraint there. The associated analytical aspects have been implemented in various contexts: asymptotically flat [45], asymptotically hyperbolic [4–6], or spatially compact [87]; see also [20, 43, 88, 157]. There exist a few other methods for constructing solutions of the constraint equations which do not require constant mean curvature: the "thin sandwich approach" of Baierlain, Sharp and Wheeler [**?**], further studied in [**?**, 19]; the gluing approach of Corvino and Schoen [49, 56, 144]; the conformal gluing technique of Joyce [97], as extended by Isenberg, Mazzeo and Pollack [90, 92]; the quasi-spherical construction of Bartnik [17, 148] and its extension due to Smith and Weinstein [151]. One can also use the implicit function theorem, or variations thereof [42, 93, 96], to construct solutions of the constraint equations for which (1.7.2) does not necessarily hold. In [20] the reader will find a presentation of alternative approaches to constructing solutions of the constraints, covering work done up to 2003.

1.7.1 The Yamabe problem

At the heart of the conformal method lies the *Yamabe problem.* From the general relativistic point of view, this correspond to special initial data where K vanishes; such initial data are called *time symmetric.* For such data (1.7.1b) becomes

$$R(g) = 2\Lambda \ . \tag{1.7.3}$$

In other words, g is a metric of constant scalar curvature.

There is a classical method, usually attributed to Yamabe [156], which allows one to construct metrics satisfying (1.7.3) by conformal deformation: given a metric $\tilde{g}$ one sets

$$g = \phi^{\frac{4}{n-2}} \tilde{g} ,$$

then (1.7.3) becomes

$$\Delta_{\widetilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi = -\frac{n-2}{2(n-1)}\Lambda\phi^{\frac{n+2}{n-2}} . \qquad (1.7.4)$$

One thus obtains a metric of constant scalar curvature 2Λ when a strictly positive solution ϕ can be found.

Equation 1.7.4 is known as the *Yamabe equation*, and the problem of finding positive solutions of this equation on compact manifolds is known as the *Yamabe problem*. The final solution, that such deformations always exist when Λ is suitably restricted (we will return to this issue shortly), has been given by Schoen [145]. Previous key contributions include [8, 153], and a comprehensive review of the problem can be found in [106]. A completely different solution has been devised by Bahri [11].

The idea is then to do something similar in general relativity, exploiting the fact that the Yamabe problem has already been solved. For this we need, first, to understand the behaviour of the vector constraint equation under conformal transformations.

Regardless of whether the manifold is compact or not, the *Yamabe number* of a metric is defined by the equation

$$Y(M,g) = \inf_{u\in C^\infty_b,\ u\not\equiv 0} \frac{\int_M (|Du|^2 + \frac{n-2}{4(n-1)}Ru^2)}{\left(\int_M u^{2n/(n-2)}\right)^{(n-2)/n}} . \qquad (1.7.5)$$

where C^∞_b denotes the space of compactly supported smooth functions. The number $Y(M,g)$ depends only upon the conformal class of g. If $Y(M,g) > 0$ we say that g is in the *positive Yamabe class*, etc. When M is compact, one can show that there exists a conformal rescaling so that $\widetilde{R}$ is positive [**?**] if and only g is in the positive Yamabe class, similarly for the zero and negative Yamabe class cases.

1.7.2 The vector constraint equation

As is made clear by the name, the *conformal method* exploits the properties of (1.7.1) under conformal transformations: consider a metric $\widetilde{g}$ related to g by a conformal rescaling:

$$\widetilde{g}_{ij} = \phi^\ell g_{ij} \quad \Longleftrightarrow \quad \widetilde{g}^{ij} = \phi^{-\ell} g^{ij} . \qquad (1.7.6)$$

This implies

$$\begin{aligned} \widetilde{\Gamma}^i{}_{jk} &= \frac{1}{2}\widetilde{g}^{im}(\partial_j\widetilde{g}_{km} + \partial_k\widetilde{g}_{jm} - \partial_m\widetilde{g}_{jk}) \\ &= \frac{1}{2}\phi^{-\ell}g^{im}(\partial_j(\phi^\ell\widetilde{g}_{km}) + \partial_k(\phi^\ell\widetilde{g}_{jm}) - \partial_m(\phi^\ell g_{jk})) \\ &= \Gamma^i{}_{jk} + \frac{\ell}{2\phi}(\delta^i_k\partial_j\phi + \delta^i_j\partial_k\phi - g_{jk}D^i\phi) , \qquad (1.7.7) \end{aligned}$$

where, as before, D denotes the covariant derivative of g.

We start by analysing what happens with (1.7.1a). Let $\widetilde{D}$ denote the covariant derivative operator of the metric $\widetilde{g}$, and consider any trace-free symmetric tensor field $\widetilde{L}^{ij}$, we have

$$\begin{aligned}\widetilde{D}_i\widetilde{L}^{ij} &= \partial_i\widetilde{L}^{ij} + \widetilde{\Gamma}^i{}_{ik}\widetilde{L}^{kj} + \widetilde{\Gamma}^j{}_{ik}\widetilde{L}^{ik} \\ &= D_i\widetilde{L}^{ij} + (\widetilde{\Gamma}^i{}_{ik} - \Gamma^i{}_{ik})\widetilde{L}^{kj} + (\widetilde{\Gamma}^j{}_{ik} - \Gamma^j{}_{ik})\widetilde{L}^{ik} \ .\end{aligned}$$

Now, from (1.7.7) we obtain

$$\begin{aligned}\widetilde{\Gamma}^i{}_{ik} &= \Gamma^i{}_{ik} + \frac{\ell}{2\phi}(\delta^i_k\partial_i\phi + \delta^i_i\partial_k\phi - g_{ik}D^i\phi) \\ &= \Gamma^i{}_{ik} + \frac{n\ell}{2\phi}\partial_k\phi \ ,\end{aligned} \tag{1.7.8}$$

and we are assuming that we are in dimension n. As $\widetilde{L}$ is traceless we obtain

$$\begin{aligned}\widetilde{D}_i\widetilde{L}^{ij} &= D_i\widetilde{L}^{ij} + \frac{n\ell}{2\phi}\partial_k\phi\widetilde{L}^{kj} + \frac{\ell}{2\phi}(\delta^j_k\partial_i\phi + \delta^j_i\partial_k\phi - \underbrace{g_{ik}D^j\phi)\widetilde{L}^{ik}}_{\sim g_{ik}\widetilde{L}^{ik}=0} \\ &= D_i\widetilde{L}^{ij} + \frac{(n+2)\ell}{2\phi}\partial_k\phi\widetilde{L}^{kj} \\ &= \phi^{-(n+2)\ell/2}D_i(\phi^{(n+2)\ell/2}\widetilde{L}^{ij}) \ .\end{aligned} \tag{1.7.9}$$

It follows that

$$\widetilde{D}_i\widetilde{L}^{ij} = 0 \quad \Longleftrightarrow \quad D_i(\phi^{(n+2)\ell/2}\widetilde{L}^{ij}) = 0 \ . \tag{1.7.10}$$

This observation leads to the following: suppose that the CMC condition (1.7.2) holds, set

$$L^{ij} := K^{ij} - \frac{\mathrm{tr}_g K}{n}g^{ij} \ . \tag{1.7.11}$$

Then L^{ij} is symmetric and trace-free whenever K^{ij} satisfies the vector constraint equation (1.7.1a). Reciprocally, let τ be any constant, and let $\widetilde{L}^{ij}$ be symmetric, trace-free, and *$\widetilde{g}$–divergence free*: by definition, this means that

$$\widetilde{D}_i\widetilde{L}^{ij} = 0 \ .$$

Set

$$L^{ij} := \phi^{(n+2)\ell/2}\widetilde{L}^{ij} \tag{1.7.12a}$$

$$K^{ij} := L^{ij} + \frac{\tau}{n}g^{ij} \ , \tag{1.7.12b}$$

then K^{ij} satisfies (1.7.1a).

More generally, assuming neither vacuum nor $d(\mathrm{tr}_g K) = 0$, with the rescaling $\widetilde{g}_{ij} = \phi^\ell g_{ij}$ and with the definitions (1.7.12) we will have

$$\begin{aligned}8\pi J^i &:= D_i(K^{ij} - \mathrm{tr}_g K g^{ij}) \\ &= D_i(\phi^{(n+2)\ell/2}\widetilde{L}^{ij}) - \frac{n-1}{n}D^j\tau \\ &= \phi^{(n+2)\ell/2}\tilde{D}_i\widetilde{L}^{ij} - \frac{n-1}{n}\phi^\ell\tilde{D}^j\tau \ .\end{aligned} \tag{1.7.13}$$

With the choice $\ell = -\frac{4}{n-2}$ which will be motivated shortly, this can also be written as the following equation for $\tilde{L}$ when τ and J^i have been given:

$$\tilde{D}_i\widetilde{L}^{ij} = 8\pi\phi^{\frac{2(n+2)}{n-2}}J^i + \frac{n-1}{n}\phi^{\frac{2n}{n-2}}\tilde{D}^j\tau \,. \tag{1.7.14}$$

1.7.3 The scalar constraint equation

To analyse the scalar constraint equation (1.7.1b) we shall use the following formula, derived in Appendix B: if $g_{ij} = \phi^\ell \tilde{g}_{ij}$, then (B.1.14) with g interchanged with $\widetilde{g}$ and ℓ changed to $-\ell$ gives

$$R(g)\phi^{-\ell} = \widetilde{R} + \frac{(n-1)\ell}{\phi}\Delta_{\widetilde{g}}\phi + \frac{(n-1)\ell\{(n-2)\ell+4\}}{4\phi^2}|d\phi|^2_{\widetilde{g}} \,, \tag{1.7.15}$$

where $\widetilde{R}$ is the scalar curvature of $\widetilde{g}$. Clearly it is convenient to choose

$$\ell = -\frac{4}{n-2} \,, \tag{1.7.16}$$

as then the last term in (1.7.7) drops out. In order to continue we use (1.7.12) to calculate

$$\begin{aligned}
|K|^2_g - (\mathrm{tr}_g K)^2 &= g_{ik}g_{jl}K^{ij}K^{kl} - \tau^2 \\
&= g_{ik}g_{jl}(L^{ij} + \frac{\tau}{n}g^{ij})(L^{kl} + \frac{\tau}{n}g^{kl}) - \tau^2 \\
&= \underbrace{g_{ik}}_{=\phi^{-\ell}g_{ik}} \; g_{jl} \; \underbrace{L^{ij}}_{=\phi^{(n/2+1)\ell}\widetilde{L}^{ij}} \; L^{kl} - \tau^2(1-\frac{1}{n}) \\
&= \phi^{n\ell}\widetilde{g}_{ik}\widetilde{g}_{jl}\widetilde{L}^{ij}\widetilde{L}^{kl} - \tau^2(1-\frac{1}{n}) \,,
\end{aligned}$$

giving thus

$$|K|^2_g - (\mathrm{tr}_g K)^2 = \phi^{n\ell}|\widetilde{L}|^2_{\widetilde{g}} - \frac{n-1}{n}\tau^2 \,. \tag{1.7.17}$$

Equations (1.7.1b), (1.7.7) and (1.7.17) with ℓ given by (1.7.16) finally yield

$$\Delta_{\widetilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi = -\tilde{\sigma}^2\phi^{(2-3n)/(n-2)} + \beta\phi^{\frac{n+2}{n-2}} \,, \tag{1.7.18}$$

where

$$\tilde{\sigma}^2 := \frac{n-2}{4(n-1)}|\widetilde{L}|^2_{\widetilde{g}} \,, \quad \beta := \left[\frac{n-2}{4n}\tau^2 - \frac{n-2}{2(n-1)}\Lambda\right] \,. \tag{1.7.19}$$

In dimension $n = 3$ this equation is known as the *Lichnerowicz equation*:

$$\boxed{\Delta_{\widetilde{g}}\phi - \frac{\widetilde{R}}{8}\phi = -\tilde{\sigma}^2\phi^{-7} + \beta\phi^5 \,.} \tag{1.7.20}$$

We note that $\tilde{\sigma}^2$ is positive, as the notation suggests, while β is a constant, non-negative if $\Lambda = 0$, or in fact if $\Lambda \leq 0$.

The strategy is now the following: let $\widetilde{g}$ be a given Riemannian metric on M, and let $\widetilde{L}^{ij}$ be any symmetric transverse $\widetilde{g}$-divergence free tensor field. We then solve (if possible) (1.7.18) for ϕ, and obtain a vacuum initial data set by calculating g using (1.7.6), and by calculating K using (1.7.12).

More generally, the energy density of matter fields is related to the geometry through the formula

$$16\pi\mu \quad := \quad R(g) - |K|^2_g + (\mathrm{tr}_g K)^2 - 2\Lambda\ . \tag{1.7.21}$$

If μ has been prescribed, this becomes an equation for ϕ

$$\begin{aligned}\Delta_{\widetilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi \quad &= \quad -\tilde{\sigma}^2\phi^{\frac{2-3n}{n-2}} + \beta\phi^{\frac{n+2}{n-2}} \\ &\quad -\frac{4(n-2)}{(n-1)}\phi^{\frac{n+2}{n-2}}\pi\mu\ .\end{aligned} \tag{1.7.22}$$

1.7.4 The vector constraint equation on compact manifolds

In order to solve the Lichnerowicz equations we need the transverse-traceless tensor (TT-tensor) field $\widetilde{L}$, and so to obtain an exhaustive construction of CMC initial data sets we have to give a prescription for constructing such tensors. It is a non-trivial fact [28] that the space of TT-tensors is always infinite dimensional in dimension larger than two.

We note that an ad-hoc example of TT-tensor on *three-dimensional non-conformally flat* manifolds is provided by the Bach tensor, see Appendix B.4. Another one is provided by the Ricci tensor on manifolds with constant scalar curvature.

A systematic prescription how to construct TT-tensors has been given by York: here one starts with an arbitrary symmetric traceless tensor field $\tilde{B}^{ij}$, which will be referred to as the *seed field*. One then writes

$$\widetilde{L}^{ij} = \tilde{B}^{ij} + \tilde{C}(Y)^{ij}\ , \tag{1.7.23}$$

where $\tilde{C}(Y)$ is the *conformal Killing operator*:

$$\tilde{C}(Y)^{ij} := \widetilde{D}^i Y^j + \widetilde{D}^j Y^i - \frac{2}{n}\widetilde{D}_k Y^k \widetilde{g}^{ij}\ . \tag{1.7.24}$$

The requirement that $\widetilde{L}^{ij}$ be divergence free becomes then an equation for the vector field Y:

$$\widetilde{D}_i\widetilde{L}^{ij} = 0 \quad \Longleftrightarrow \quad \widetilde{L}(Y)^j := \widetilde{D}_i(\widetilde{D}^i Y^j + \widetilde{D}^j Y^i - \frac{2}{n}\widetilde{D}_k Y^k \widetilde{g}^{ij}) = -\widetilde{D}_i\tilde{B}^{ij}\ . \tag{1.7.25}$$

While (1.7.25) looks complicated at first sight, it is rather natural: we want to produce transverse traceless tensors by solving an elliptic differential equation. Since the condition of being divergence-free is already a first order equation, and it is not elliptic, then the lowest possible order of such an equation will be two. Now, the divergence operation turns two-contravariant tensor fields to vector fields, so the most straightforward way of ensuring ellipticity is to seek an equation for a vector field. The simplest object that we obtain by differentiating a vector field is the tensor field $\widetilde{D}^i Y^j$; in order to achieve the desired symmetries we need to symmetrise and remove the trace, which leads to the conformal Killing operator (1.7.24).

The operator L defined in (1.7.25) is known as the *conformal vector Laplacian.* Equation (1.7.25) is a second order linear partial differential equation for Y, the solvability of which can be easily analysed. In this section we shall consider spatially compact manifolds M. We will give an existence proof for (1.7.25):

THEOREM 1.7.1 *For any smooth symmetric traceless tensor field $\tilde{B}^{ij}$ there exists a smooth vector field Y such that (1.7.25) holds.*

PROOF: Recall that a *conformal Killing vector* for the metric $\widetilde{g}$ is a nontrivial solutions of the equation $\tilde{C}(Y) = 0$. When $(M, \widetilde{g})$ does not admit any conformal Killing vectors, Theorem 1.7.1 follows immediately from Theorem 1.7.6 below together with (1.7.35) and (1.7.37), because then the equation

$$\widetilde{L}(Y) = Z$$

has a solution for any Z.

When conformal Killing vectors exist, the image of $\widetilde{L}$ is the L^2-orthogonal of the kernel of $\widetilde{L}^\dagger$. Since $\widetilde{L}$ is formally self-adjoint, the image of $\widetilde{L}$ is orthogonal to the space of conformal Killing vectors. But the right-hand side of (1.7.25) is orthogonal to that last space: indeed, if Z is a conformal Killing vector and $\tilde{B}^{ij}$ is a symmetric traceless tensor field, then integration by parts gives

$$\begin{aligned}
\int_M Z_i \widetilde{D}_j \tilde{B}^{ij} &= -\int_M \widetilde{D}_j Z_i \tilde{B}^{ij} \\
&= -\frac{1}{2}\int_M (\widetilde{D}_j Z_i + \widetilde{D}_i Z_j)\tilde{B}^{ij} \qquad (\tilde{B} \text{ is symmetric}) \\
&= -\frac{1}{2}\int_M \underbrace{(\widetilde{D}_j Z_i + \widetilde{D}_i Z_j - \frac{1}{2}\widetilde{D}_k Z^k \widetilde{g}_{ij})}_{0}\tilde{B}^{ij} \qquad (\tilde{B} \text{ is trace-free}) \\
&= 0\,. \qquad (1.7.26)
\end{aligned}$$

This shows that $-\widetilde{D}_j \tilde{B}^{ij}$ lies in the image of $\widetilde{L}$, and so there exist many solutions of (1.7.25). (Note that the non-uniqueness does not change $\tilde{L}^{ij}$, as defined in (1.7.23).) □

A property essentially equivalent to Theorem 1.7.1 is the existence of the *York splitting*, also known in the mathematical literature as the *Berger-Ebin splitting*:

THEOREM 1.7.2 *On any compact Riemannian manifold (M, g) the space of symmetric tensors, say $\Gamma S^2 M$, splits L^2-orthogonally as*

$$\Gamma S^2 M = C^\infty g \oplus TT \oplus \mathrm{Im} C\ ,$$

where $C^\infty g$ are tensors proportional to the metric, TT denotes the space of transverse traceless tensors, and $\mathrm{Im}C$ is the image of the conformal Killing operator defined in (1.7.24).

PROOF: : Given any symmetric two-covariant tensor field A let ψ denote the trace of A dived by n, set

$$B_{ij} = A_{ij} - \psi g_{ij} \; .$$

Then B_{ij} is symmetric and traceless. Similarly to (1.7.25), we let Y be any solution of the equation

$$L(Y)^i = D_j \tilde{B}^{ij} \; .$$

Here, of course, $C(Y)^{ij} := D^i Y^j + D^j Y^i - \frac{2}{n} D_k Y^k g^{ij}$ and $L(Y)^i = D_i C(Y)^{ij}$. Then $B_{ij} - C(Y)_{ij}$ is transverse and traceless, and we have indeed

$$A_{ij} = \underbrace{\psi g_{ij}}_{\in C^\infty \times g} + \underbrace{B_{ij} - C(Y)_{ij}}_{\in TT} + \underbrace{C(Y)_{ij}}_{\in \mathrm{Im} C} \; .$$

The L^2-orthogonality of the factors is easily verified; compare (1.7.26). □

1.7.5 Some linear elliptic theory

The main ingredients of the existence proof which we will present shortly are the following:

1. *Function spaces*: one uses the spaces H_k, $k \in \mathbb{N}$, defined as the completion of the space of smooth tensor fields on M with respect to the norm

$$\|u\|_k := \sqrt{\sum_{0 \le \ell \le k} \int_M |D^\ell u|^2 d\mu} \; , \tag{1.7.27}$$

where $D^\ell u$ is the tensor of ℓ-th covariant derivatives of u with respect to some covariant derivative operator D. For compact manifolds[2] this space is identical with that of fields in L^2 such that their distributional derivatives of order less than or equal to k are also in L^2. Again for compact manifolds, different choices of measure $d\mu$ (as long as it remains absolutely continuous with respect to the coordinate one), of the tensor norm $|\cdot|$, or of the connection D, lead to the same space, with equivalent norm.

Recall that if $u \in L^2$ then $\partial_i u = \rho_i$ in a distributional sense if for every smooth compactly supported vector field we have

$$\int_M X^i \rho_i = - \int_M D_i X^i u \; .$$

More generally, let A be a linear differential operator of order m and let A^t be its *formal* L^2 adjoint, which is the operator obtained by differentiating by parts:

$$\int_M \langle u, L^\dagger v \rangle := \int_M \langle Lu, v \rangle \; , \quad u, v \in C^m_c;$$

the above formula defines $L^\dagger$ uniquely if it holds for all u, v in the space C^m_c of C^m compactly supported fields. (Incidentally, the reader will note by

[2] For non-compact manifolds this is not always the case, compare [9].

comparing the last two equations that the formal adjoint of the derivative operator is minus the divergence operator.) Then, for $u \in L^1_{\rm loc}$ (this is the space of measurable fields u which are Lebesgue-integrable on any compact subset of the manifold), the distributional equation $Lu = \rho$ is said to hold if for all smooth compactly supported v's we have

$$\int_M \langle u, L^\dagger v\rangle = \int_M \langle \rho, v\rangle \ .$$

One sometimes talks about *weak solutions* rather than distributional ones.

The spaces H_k are Hilbert spaces with the obvious scalar product:

$$\langle u, v\rangle_k = \sum_{0\le \ell\le k} \int_M \langle D^\ell u, D^\ell v\rangle d\mu \ .$$

The Sobolev embedding theorem [10] asserts that H_k functions are, locally, of $C^{k'}$ differentiability class, where k' is the largest integer satisfying

$$k' < k - n/2 \ . \tag{1.7.28}$$

On a compact manifold the result is true globally,

$$H_k \subset C^{k'} \ , \tag{1.7.29}$$

with the inclusion map being continuous:

$$\|u\|_{C^{k'}} \le C\|u\|_{H_k} \ . \tag{1.7.30}$$

2. *Orthogonal complements in Hilbert spaces:* Let H be a Hilbert space, and let E be a *closed linear subspace* of H. Then (see, *e.g.*, [154]) we have the direct sum

$$H = E \oplus E^\perp \ . \tag{1.7.31}$$

This result is sometimes called *the projection theorem.*

3. *Rellich-Kondrashov compactness:* we have the obvious inclusion

$$H_k \subset H_{k'} \ \text{ if } \ k \ge k' \ .$$

The *Rellich-Kondrashov theorem* (see, *e.g.*, [1, 10, 78, 101]) asserts that, on compact manifolds, this inclusion is *compact.* Equivalently,[3] if u_n is any sequence satisfying $\|u_n\|_k \le C$, and if $k' < k$, then there exists a subsequence u_{n_i} and $u_\infty \in H_k$ such that u_{n_i} converges to u_∞ in $H_{k'}$ topology as i tends to infinity.

[3] In this statement we have also made use of the *Tichonov-Alaoglu* theorem, which asserts that bounded sets in Hilbert spaces are weakly compact; *cf., e.g.* [154].

4. *Elliptic regularity:* If $Y \in L^2$ satisfies $LY \in H_k$ in a distributional sense, with L — an elliptic operator of order m with smooth coefficients, then $Y \in H_{k+m}$, and Y satisfies the equation in the classical sense. Further, on compact manifolds for every k there exists a constant C_k such that

$$\|Y\|_{k+m} \le C_k(\|LY\|_k + \|Y\|_0) . \tag{1.7.32}$$

Our aim is to show that solvability of (1.7.25) can be easily studied using the above basic facts. We start by verifying ellipticity of L. Recall that the *symbol* σ of a linear partial differential operator L of the form

$$L = \sum_{0\le \ell \le m} a^{i_1\dots i_\ell} D_{i_1}\dots D_{i_\ell} ,$$

where the $a^{i_1\dots i_\ell}$'s are linear maps from fibers of a bundle E to fibers of a bundle F, is defined as the map

$$T^*M \ni p \mapsto \sigma(p) := a^{i_1\dots i_m} p_{i_1}\dots p_{i_m} .$$

Thus, every derivative D_i is replaced by p_i, and all terms other than the top order ones are ignored. An operator is said to be *elliptic*[4] if the symbol is an isomorphism of fibers for all $p \neq 0$. In our case (1.7.25) the operator L acts on vector fields and produces vector fields, with

$$TM \ni Y \to \sigma(p)(Y) = p_i(p^iY^j + p^jY^i - \frac{2}{n}p_kY^k\tilde g^{ij})\partial_j \in TM . \tag{1.7.33}$$

(The indices on p^i have been raised with the metric $\tilde g$.) To prove bijectivity of $\sigma(p)$, $p \neq 0$, it suffices to verify that $\sigma(p)$ has trivial kernel. Assuming $\sigma(p)(Y) = 0$, a contraction with p_j gives

$$p_jp_i(p^iY^j + p^jY^i - \frac{2}{n}p_kY^k\tilde g^{ij}) = |p|^2p_jY^j(2 - \frac{2}{n}) = 0 ,$$

hence $p_jY^j = 0$ for $n > 1$ since $p \neq 0$. Contracting instead with Y_j and using the last equality we obtain

$$Y_jp_i(p^iY^j + p^jY^i - \frac{2}{n}p_kY^k\tilde g^{ij}) = |p|^2|Y|^2 = 0 ,$$

and $\sigma(p)$ has no kernel, as desired.

To gain some more insight into the conformal vector Laplacian L let us calculate its formal L^2–adjoint: let thus X and Y be smooth, or C^2, we write

$$\begin{aligned}\int_M X_iL(Y)^i d\mu_{\tilde g} &= \int_M X_i\tilde D_j(\tilde D^iY^j + \tilde D^jY^i - \frac{2}{n}\tilde g^{ij}\tilde D_kY^k)\, d\mu_{\tilde g} \\ &= -\int_M \tilde D_jX_i(\tilde D^iY^j + \tilde D^jY^i - \frac{2}{n}\tilde g^{ij}\tilde D_kY^k)\, d\mu_{\tilde g}\end{aligned}$$

[4] See [2, 127] for more general notions of ellipticity.

$$
\begin{aligned}
&= -\frac{1}{2}\int_M (\widetilde{D}_j X_i + \widetilde{D}_j X_i)\underbrace{(\widetilde{D}^i Y^j + \widetilde{D}^j Y^i - \frac{2}{n}\widetilde{g}^{ij}\widetilde{D}_k Y^k)}_{\text{symmetric in } i \text{ and } j}\, d\mu_{\widetilde{g}} \\
&= -\frac{1}{2}\int_M (\widetilde{D}_j X_i + \widetilde{D}_j X_i - \frac{2}{n}\widetilde{D}_k X^k \widetilde{g}_{ij})\underbrace{(\widetilde{D}^i Y^j + \widetilde{D}^j Y^i - \frac{2}{n}\widetilde{g}^{ij}\widetilde{D}_k Y^k)}_{\text{trace free}}\, d\mu_{\widetilde{g}} \\
&= -\int_M (\widetilde{D}_j X_i + \widetilde{D}_j X_i - \frac{2}{n}\widetilde{D}_k X^k \widetilde{g}_{ij})\widetilde{D}^i Y^j \, d\mu_{\widetilde{g}} \\
&= \int_M \widetilde{D}^i(\widetilde{D}_j X_i + \widetilde{D}_j X_i - \frac{2}{n}\widetilde{D}_k X^k \widetilde{g}_{ij}) Y^j \, d\mu_{\widetilde{g}} \\
&= \int_M L(X)^j Y_j \, d\mu_{\widetilde{g}} \, . \qquad (1.7.34)
\end{aligned}
$$

Recall that the *formal adjoint* $L^\dagger$ of L is defined by integration by parts:

$$\int \langle u, Lv\rangle = \int \langle L^\dagger u, v\rangle$$

for all smooth compactly supported fields u, v. (Note that the definition of a self-adjoint operator further requires an equality of domains, an issue which is, fortunately, completely ignored in the formal definition.) We have thus shown that the conformal vector Laplacian is *formally self adjoint*:

$$L^\dagger = L \, . \qquad (1.7.35)$$

We further note that the fourth line in (1.7.34) implies

$$\int_M Y_i L(Y)^i = -\frac{1}{2}\int_M |C(Y)|^2 \, , \qquad (1.7.36)$$

in particular if Y is C^2 then

$$L(Y) = 0 \quad \Longleftrightarrow \quad C(Y) = 0 \, . \qquad (1.7.37)$$

This implies that Riemannian manifolds for which L has a non-trivial kernel are very special.

REMARK 1.7.3 Solutions of the equation $C(Y) = 0$ are called *conformal Killing vectors*. The existence of non-trivial conformal Killing vectors implies the existence of conformal isometries of (M, g). A famous theorem of Lelong-Ferrand – Obata [107, 130] (compare [105]) shows that, on compact manifolds in dimensions greater than or equal to three, there exists a conformal rescaling such that Y is a Killing vector, except if (M, g) is conformally isometric to S^n with a round metric. In the former case (the conformally rescaled) (M, g) has a non-trivial isometry group, which imposes restrictions on the topology of M, and forces g to be very special. For instance, the existence of non-trivial Lie group of isometries of a compact manifold implies that M admits an S^1 action, which is a serious topological restriction, and in fact is not possible for "most" topologies (see, *e.g.*, [64, 65], and also [66] and references therein for an analysis in dimension four). It is also true that even if M admits S^1 actions, then there exists an open and dense set of metrics, in a $C^{k(n)}$ topology, or in a $H^{k'(n)}$ topology, with appropriate $k(n)$, $k'(n)$ [23], for which no nontrivial solutions of the over-determined system of equations $C(Y) = 0$ exist.

In order to continue we shall need a somewhat stronger version of (1.7.32):

PROPOSITION 1.7.4 *Let L be an elliptic operator of order m on a compact manifold. If there are no non-trivial smooth solutions of the equation $L(Y) = 0$, then (1.7.32) can be strengthened to*

$$\|Y\|_{k+m} \leq C'_k \|L(Y)\|_k \ . \tag{1.7.38}$$

REMARK 1.7.5 Equation (1.7.38) implies that L has trivial kernel, which shows that the condition on the kernel is necessary.

PROOF: Suppose that the result does not hold, then for every $n \in \mathbb{N}$ there exists $Y_n \in H_{k+m}$ such that

$$\|Y_n\|_{k+m} \geq n\|L(Y_n)\|_k \ . \tag{1.7.39}$$

Multiplying Y_n by an appropriate constant if necessary we can suppose that

$$\|Y\|_{L^2} = 1 \ . \tag{1.7.40}$$

The basic elliptic inequality (1.7.32) gives

$$\|Y_n\|_{k+m} \leq C_2(\|LY_n\|_k + \|Y_n\|_0) \leq \frac{C_2}{n}\|Y_n\|_{k+m} + C_2 \ ,$$

so that for n such that $C_2/n \leq 1/2$ we obtain

$$\|Y_n\|_{k+m} \leq 2C_2 \ .$$

It follows that Y_n is bounded in H_{k+m}; further (1.7.39) gives

$$\|L(Y_n)\|_k \leq \frac{2C_2}{n} \ . \tag{1.7.41}$$

By the Rellich-Kondrashov compactness we can extract a subsequence, still denoted by Y_n, such that Y_n converges in L^2 to $Y_* \in H_{k+m}$. Continuity of the norm together with L^2 convergence implies that

$$\|Y_*\|_{L^2} = 1 \ , \tag{1.7.42}$$

so that $Y_* \neq 0$. One would like to conclude from (1.7.41) that $L(Y_*) = 0$, but that is not completely clear because we do not know whether or not

$$LY_* = \lim_{n\to\infty} LY_n \ .$$

Instead we write the distributional equation: for every smooth X we have

$$\int_M \langle L(Y_n), X\rangle = \int_M \langle Y_n, L^\dagger(X)\rangle \ .$$

Now, $L(Y_n)$ tends to zero in L^2 by (1.7.41), and Y_n tends to Y_* in L^2, so that passing to the limit we obtain

$$0 = \int_M \langle Y_*, L^\dagger(X)\rangle \ .$$

It follows that Y_* satisfies $L(Y_*) = 0$ in a distributional sense. Elliptic regularity implies that Y_* is a smooth solution of $LY_* = 0$, it is non-trivial by (1.7.42), a contradiction. □

We are ready to prove now:

THEOREM 1.7.6 *Let L be any elliptic partial differential operator of order m on a compact manifold and suppose that the equations $Lu = 0$, $L^\dagger v = 0$ have no non-trivial smooth solutions, where $L^\dagger$ is the formal adjoint of L. Then for any $k \geq 0$ the map*

$$L : H_{k+m} \to H_k$$

is an isomorphism.

PROOF: An element of the kernel is necessarily smooth by elliptic regularity, it remains thus to show surjectivity. We start by showing that the image of L is closed: let Z_n be a Cauchy sequence in $\operatorname{Im} L$, then there exists $Z_\infty \in L^2$ and $Y_n \in H_{k+m}$ such that

$$LY_n = Z_n \stackrel{L^2}{\to} Z_\infty \ .$$

Applying (1.7.38) to $Y_n - Y_\ell$ we find that Y_n is Cauchy in H_{k+m}, therefore converges in H_{k+m} to some element $Y_\infty \in H_{k+m}$. By continuity of L the sequence LY_n converges to LY_∞ in L^2, hence $Z_\infty = LY_\infty$, as desired.

Consider, first, the case $k = 0$. By the orthogonal decomposition theorem we have now

$$L^2 = \operatorname{Im} L \oplus (\operatorname{Im} L)^\perp \ ,$$

and if we show that $(\operatorname{Im} L)^\perp = \{0\}$ we are done. Let, thus, $Z \in (\operatorname{Im} L)^\perp$, this means that

$$\int_M \langle Z, L(Y) \rangle = 0 \qquad (1.7.43)$$

for all $Y \in H_{m+2}$. In particular (1.7.43) holds for all smooth Y, which implies that $L^\dagger(Z) = 0$ in a distributional sense. Now, the symbol of $L^\dagger$ is the transpose of the symbol of L, which shows that $L^\dagger$ is also elliptic. We can thus use elliptic regularity to conclude that Z is smooth, and $Z = 0$ follows.

The result in L^2 together with elliptic regularity immediately imply the result in H_k. □

1.7.6 The scalar constraint equation on compact manifolds, $\tau^2 \geq \frac{2n}{(n-1)}\Lambda$

Theorem 1.7.6, together with Equation (1.7.37) and Remark 1.7.3, gives a reasonably complete description of the solvability of (1.7.25). We simply note that if $\tilde{B}^{ij}$ there is smooth, then the associated solution will be smooth by elliptic regularity. To finish the presentation of the conformal method we need to address the question of existence of solutions of the Lichnerowicz equation (1.7.18).

A complete description can be obtained when the constant β defined in (1.7.19) satisfies

$$\beta := \left[\frac{n-2}{4n}\tau^2 - \frac{n-2}{2(n-1)}\Lambda\right] \geq 0 \ , \qquad (1.7.44)$$

this will certainly be the case if $\Lambda \leq 0$. In this case, to emphasise positivity we will write

$$\frac{n-2}{4n}\tau_\Lambda^2$$

for β, thus rewriting (1.7.18) as

$$\Delta_{\tilde g}\phi - \frac{n-2}{4(n-1)}\widetilde R\phi = -\tilde\sigma^2\phi^{(2-3n)/(n-2)} + \frac{n-2}{4n}\tau_\Lambda^2\phi^{\frac{n+2}{n-2}} \,. \qquad (1.7.45)$$

As already pointed out in Section 1.7.1, the case $\sigma = 0$ corresponds to the so-called *Yamabe* equation; in this case solutions of (1.7.45) produce metrics with constant scalar curvature $-(n-1)\tau_\Lambda^2$. We will take it for granted that one can first deform the metric conformally so that $\widetilde R$ is constant, and we will assume that this has been done. It should be recognised that making use of the solution of the Yamabe problem sweeps the real difficulties under the carpet. Nevertheless, there remains some analysis to do even after the Yamabe part of the problem has been solved.

In what follows we will assume smoothness of all objects involved. More recently, these equations have been studied with metrics of low differentiability [34, 115]; this was motivated in part by work on the evolution problem for "rough initial data" [98–100, 152]. Boundary value problems for the constraint equations, with nonlinear boundary conditions motivated by black holes, were considered in [59, 116].

In order to provide a complete answer to the question of solvability of (1.7.45), as first done by Isenberg [87], we start by showing that (1.7.45) has no solutions in several cases: For this, suppose that there exists a solution, and integrate (1.7.45) over M:

$$\int_M \left(\frac{n-2}{4(n-1)}\widetilde R - \tilde\sigma^2\phi^{(2-3n)/(n-2)} + \frac{n-2}{4n}\tau_\Lambda^2\phi^{\frac{n+2}{n-2}}\right) = 0 \,.$$

Since we want ϕ to be positive, there are obvious obstructions for this equation to hold, and hence for existence of positive solutions: for example, if $\tilde\sigma^2 \equiv 0$ and $\tau_\Lambda^2 = 0$ then there can be a positive solution only if $\widetilde R$ vanishes (and then ϕ is necessarily constant, e.g. by an appropriate version of the maximum principle). Analysing similarly other possibilities one finds:

Proposition 1.7.7 *Suppose that*

1. $\tilde\sigma^2 \equiv 0 \equiv \tau_\Lambda^2$, *but* $\tilde R \neq 0$;
2. $\tilde\sigma^2 \equiv 0$, $\tau_\Lambda^2 \neq 0$ *but* $\tilde R \geq 0$;
3. $\tau_\Lambda^2 \equiv 0$, $\tilde\sigma^2 \not\equiv 0$, *but* $\tilde R \leq 0$.

Then (1.7.45) has no positive solutions.

We emphasize that the non-existence result is *not* a failure of the conformal method to produce solutions, but a *no-go* result; we will return to this issue in Proposition 1.7.16 below.

It turns out that there exist positive solutions for all other cases. This will be proved using the *monotone iteration scheme*, which we are going to describe now. For completeness we start by proving a simple version of the *maximum principle*:

PROPOSITION 1.7.8 *Let (M,g) be compact, suppose that $c<0$ and let $u\in C^2(M)$. If*

$$\Delta u + cu \geq 0\ , \tag{1.7.46}$$

then $u\leq 0$. If equality in (1.7.46) holds then $u\equiv 0$.

PROOF: Suppose that u has a strictly positive maximum at p. In local coordinates around p we then have

$$g^{ij}\partial_i\partial_j u - g^{ij}\Gamma^k{}_{ij}\partial_k u \geq -cu\ .$$

The second term on the left-hand-side vanishes at p because ∂u vanishes at p, the first term is non-positive because at a maximum the matrix of second partial derivatives is non-positive definite. On the other hand the right-hand-side is strictly positive, which gives a contradiction. If equality holds in (1.7.46) then both u and minus u are non-positive, hence the result. □

Consider, now, the operator

$$L = \Delta_{\widetilde{g}} + c$$

for some $c<0$. The symbol of L reads

$$\sigma_L(p) = g^{ij}p_ip_j \neq 0 \ \text{ if } \ p\neq 0\ ,$$

which shows that L is elliptic. It is well-known that $\Delta_{\widetilde{g}}$ is formally self-adjoint (with respect to the measure $d\mu_{\widetilde{g}}$), and Proposition 1.7.8 allows us to apply Theorem 1.7.6 to conclude existence of H_{k+2} solutions of the equation

$$Lu = \rho \tag{1.7.47}$$

for any $\rho\in H_k$; u is smooth if ρ and the metric are.

Returning to the Lichnerowicz equation (1.7.45), let us rewrite this equation in the form

$$\Delta_{\widetilde{g}}\phi = F(\phi,x)\ . \tag{1.7.48}$$

A C^2 function ϕ_+ is called a *super-solution* of (1.7.48) if

$$\Delta_{\widetilde{g}}\phi_+ \leq F(\phi_+,x)\ . \tag{1.7.49}$$

Similarly a C^2 function ϕ_- is called a *sub-solution* of (1.7.48) if

$$\Delta_{\widetilde{g}}\phi_- \geq F(\phi_-,x)\ . \tag{1.7.50}$$

A solution is both a sub-solution and a super-solution. This shows that a necessary condition for existence of solutions is the existence of sub- and super-solutions. It turns out that this condition is also sufficient, modulo an obvious inequality between ϕ_- and ϕ_+:

THEOREM 1.7.9 *Suppose that (1.7.48) admits a sub-solution ϕ_- and a super-solution ϕ_+ satisfying*

$$\phi_- \leq \phi_+\ .$$

If F is differentiable in ϕ, then there exists a C^2 solution ϕ of (1.7.48) such that

$$\phi_- \leq \phi \leq \phi_+\ .$$

(ϕ is smooth if F is.)

PROOF: The argument is known as the *monotone iteration scheme*, or the *method of sub- and super-solutions.* We set

$$\phi_0 = \phi_+ \ ,$$

and our aim is to construct a sequence of functions such that

$$\phi_- \leq \phi_n \leq \phi_+ \ , \tag{1.7.51a}$$

$$\phi_{n+1} \leq \phi_n \ . \tag{1.7.51b}$$

We start by chosing c to be a positive constant large enough so that the function

$$\phi \to F_c(\phi, x) := F(\phi, x) - c\phi$$

is monotone decreasing for $\phi_- \leq \phi \leq \phi_+$. This can clearly be done on a compact manifold. By what has been said we can solve the equation

$$(\Delta_{\widetilde{g}} - c)\phi_{n+1} = F_c(\phi_n, x) \ .$$

Clearly (1.7.51a) holds with $n = 0$. Suppose that (1.7.51a) holds for some n, then

$$\begin{aligned} (\Delta_{\widetilde{g}} - c)(\phi_{n+1} - \phi_+) &= F_c(\phi_n, x) - \underbrace{\Delta_{\widetilde{g}}\phi_+}_{\leq F(\phi_+, x)} - c\phi_+ \\ &\geq F_c(\phi_n, x) - F_c(\phi_+, x) \geq 0 \ , \end{aligned}$$

by monotonicity of F_c. The maximum principle gives

$$\phi_{n+1} \leq \phi_+ \ ,$$

and induction establishes the second inequality in (1.7.51a). Similarly we have

$$\begin{aligned} (\Delta_{\widetilde{g}} - c)(\phi_- - \phi_{n+1}) &= \underbrace{\Delta_{\widetilde{g}}\phi_-}_{\geq F(\phi_-, x)} - c\phi_- - F_c(\phi_n, x) \\ &\geq F_c(\phi_-, x) - F_c(\phi_n, x) \geq 0 \ , \end{aligned}$$

and (1.7.51a) is established. Next, we note that (1.7.51a) implies (1.7.51b) with $n = 0$. To continue the induction, suppose that (1.7.51b) holds for some $n \geq 0$, then

$$(\Delta_{\widetilde{g}} - c)(\phi_{n+2} - \phi_{n+1}) = F_c(\phi_{n+1}, x) - -F_c(\phi_n, x) \geq 0 \ ,$$

again by monotonicity of F_c, and (1.7.51b) is proved.

Since ϕ_n is monotone decreasing and bounded there exists ϕ such that ϕ_n tends pointwise to ϕ as n tends to infinity. Continuity of F gives

$$F_n := F(\phi_n, x) \to F_\infty = F(\phi, x) \ ,$$

again pointwise. By the Lebesgue dominated theorem F_n converges to F_∞ in L^2, and the elliptic inequality (1.7.32) gives

$$\|\phi_n - \phi_m\|_{H^2} \leq C_2(\|(\Delta_{\widetilde{g}} - c)(\phi_n - \phi_m)\|_{L^2} + \|\phi_n - \phi_m\|_{L^2} \ .$$

Completeness of H^2 implies that there exists $\phi_\infty \in H^2$ such that $\phi_n \to \phi_\infty$ in H^2. Recall that from any sequence converging in L^2 we can extract a subsequence converging pointwise almost everywhere, which shows that $\phi = \phi_\infty$ almost everywhere, hence $\phi \in H^2$. Continuity of $\Delta_{\widetilde{g}} + c$ on H^2 shows that

$$(\Delta_{\widetilde{g}} - c)\phi = \lim_{n\to\infty} (\Delta_{\widetilde{g}} - c)\phi_n = F_c(\phi, x) = F(\phi, x) - c\phi \ ,$$

so that ϕ satisfies the equation, as desired. The remaining claims follow from elliptic regularity theory. □

In order to apply Theorem 1.7.9 to the Lichnerowicz equation (1.7.45) we need appropriate sub- and super-solutions. The simplest guess is to use constants, and we start by exploring this possibility. Setting $\phi_- = \epsilon$ for some small constant $\epsilon > 0$, we need

$$0 = \Delta_{\widetilde{g}}\epsilon \geq F(\epsilon, x) \equiv \frac{n-2}{4(n-1)}\widetilde{R}\epsilon - \tilde{\sigma}^2\epsilon^{(2-3n)/(n-2)} + \frac{n-2}{4n}\tau_\Lambda^2\epsilon^{\frac{n+2}{n-2}} \qquad (1.7.52)$$

for ϵ small enough. Since $2-3n$ is negative and $\frac{n+2}{n-2}$ is larger than one, we find:

Lemma 1.7.10 *A sufficiently small positive constant is a subsolution of (1.7.45) if*

1. $\widetilde{R} < 0$, *or if*
2. $\tilde{\sigma}^2 > 0$.

Next, we set $\phi_+ = M$, with M a large constant, and we need to check that

$$0 \leq \frac{n-2}{4(n-1)}\widetilde{R}M - \tilde{\sigma}^2 M^{(2-3n)/(n-2)} + \frac{n-2}{4n}\tau_\Lambda^2 M^{\frac{n+2}{n-2}} \ . \qquad (1.7.53)$$

We see that:

Lemma 1.7.11 *A sufficiently large positive constant is a supersolution of (1.7.45) if*

1. $\widetilde{R} > 0$, *or if*
2. $\tau_\Lambda^2 > 0$.

As an immediate Corollary of the two Lemmata and of Theorem 1.7.9 one has:

Corollary 1.7.12 *The Lichnerowicz equation can always be solved if* $\widetilde{R}$ *is strictly negative and* $\tau_\Lambda \neq 0$.

Before proceeding further it is convenient to classify the metrics on M as follows: we shall say that $g \in \mathscr{Y}^+$ if g can be conformally deformed to achieve positive scalar curvature. We shall say that $g \in \mathscr{Y}^0$ if g can be conformally rescaled to achieve zero scalar curvature but $g \notin \mathscr{Y}^+$. Finally, we let $\mathscr{Y}^-$ be the collection of the remaining metrics. It is known that all classes are non-empty, and that every metric belongs to precisely one of the classes.

One then has the following result of Isenberg [87]:

THEOREM 1.7.13 *The following table summarizes whether or not the Lichnerowicz equation (1.7.45) admits a positive solution:*

	$\tilde\sigma^2 \equiv 0,\ \tau_\Lambda = 0$	$\tilde\sigma^2 \equiv 0,\ \tau_\Lambda \neq 0$	$\tilde\sigma^2 \not\equiv 0,\ \tau_\Lambda = 0$	$\tilde\sigma^2 \not\equiv 0,\ \tau_\Lambda \neq 0$
$\widetilde{g} \in \mathscr{Y}^+$	*no*	*no*	*yes*	*yes*
$\widetilde{g} \in \mathscr{Y}^0$	*yes*	*no*	*no*	*yes*
$\widetilde{g} \in \mathscr{Y}^-$	*no*	*yes*	*no*	*yes*

For initial data in the class $(\mathscr{Y}^0, \sigma \equiv 0, \tau_\Lambda = 0)$ *all solutions are constants, and any positive constant is a solution. In all other cases the solutions are unique.*

PROOF: All the "no" entries are covered by Proposition 1.7.7. The "yes" in the first column follows from the fact that constants are (the only) solutions in this case.

To cover the remaining "yes" entries, let us number the rows and columns of the table as in a matrix T_{ij}. Then T_{32} and T_{34} are the contents of Corollary 1.7.12.

In the positive Yamabe class, Lemma 1.7.11 shows that a sufficiently large constant provides a supersolution. A small constant provides a subsolution if $\tilde\sigma^2$ has no zeros; this establishes T_{13} and T_{14} for strictly positive $\tilde\sigma^2$. However, it could happen that $\tilde\sigma^2$ has zeros. To cover this case, as well as the zero-Yamabe-class case T_{24}, we use a mixture of an unpublished argument of E. Hebey [82] and of that in [117]. Similarly to several claims above, this applies to the following general setting: Let h, a, and f be smooth functions on a compact Riemannian manifold M, with $h \geq 0$, $a \geq 0$ and $f \geq 0$. Consider the equation

$$\Delta_{\tilde g} u - hu = fu^\alpha - au^{-\beta} , \tag{1.7.54}$$

with $\alpha > 1$ and $\beta > 0$. We further require $f + h \not\equiv 0$ and $a \not\equiv 0$. (All those hypotheses are satisfied in T_{13}, T_{14}, and T_{24}.) Then there exists a function u_1 such that

$$\Delta_{\tilde g} u_1 - (h + f)u_1 = -a .$$

The function u_1 is strictly positive by the maximum principle. For $t > 0$ sufficiently small the function $u_t = tu_1$ is a subsolution of (1.7.54): indeed, from $ta \leq at^{-\beta}u_1^{-\beta}$ and $ftu_1 \geq ft^\alpha u_1^\alpha$ for t small enough we conclude that

$$\Delta_{\tilde g} u_t - hu_t = -ta + tu_1 f \geq -at^{-\beta}u_1^{-\beta} + ft^\alpha u_1^\alpha .$$

The existence of a solution follows again from Theorem 1.7.9.

Uniqueness in all $\tilde{R} \geq 0$ cases, except T_{21}, follows from the fact that the function $\phi \mapsto F(\phi, x)$, defined in (1.7.52), is monotonously increasing for non-negative $\tilde{R}$: indeed, let ϕ_1 and ϕ_2 be two solutions of (1.7.48), then

$$\Delta_{\tilde{g}}(\phi_2 - \phi_1) + (\underbrace{-\int_{\phi_1}^{\phi_2} \partial_\phi F(\phi, x) d\phi}_{=:c})(\phi_2 - \phi_1) = 0 \; .$$

It follows from the monotonicity properties of F that $c \leq 0$. A version of the maximum principle, slightly stronger than the one proved in Proposition 1.7.8, gives $\phi_1 = \phi_2$ whenever the function c is *not* identically zero.

To prove uniqueness when $\tilde{R} < 0$, suppose that there exist two distinct solutions ϕ_a, $a = 1, 2$; exchanging the ϕ_a's if necessary we can without loss of generality assume that there exist points such that $\phi_2 > \phi_1$. By construction, the scalar curvature R of the metric $g := \phi_2^{\frac{4}{n-2}} \tilde{g}$ satisfies

$$\frac{n-2}{4(n-1)} R = \tilde{\sigma}^2 - \frac{n-2}{4n} \tau_\Lambda^2 \; . \tag{1.7.55}$$

Because the whole construction is conformally covariant, the function

$$\phi := \frac{\phi_1}{\phi_2}$$

satisfies again (1.7.45) with respect to the metric g:

$$\Delta_g \phi - \frac{n-2}{4(n-1)} R\phi = -\tilde{\sigma}^2 \phi^{(2-3n)/(n-2)} + \frac{n-2}{4n} \tau_\Lambda^2 \phi^{\frac{n+2}{n-2}} \; . \tag{1.7.56}$$

In view of (1.7.55), this can be rewritten as

$$\Delta_g \phi = -\tilde{\sigma}^2 (\phi^{(2-3n)/(n-2)} - \phi) + \frac{n-2}{4n} \tau_\Lambda^2 (\phi^{\frac{n+2}{n-2}} - \phi) \; . \tag{1.7.57}$$

By choice, the minimum value of ϕ, say a, is strictly smaller than one. At the point where the minimum is attained we obtain

$$0 \leq \Delta_g \phi = \underbrace{-\tilde{\sigma}^2 (a^{-\frac{3n-2}{n-2}} - a)}_{I} + \underbrace{\frac{n-2}{4n} \tau_\Lambda^2 (a^{\frac{n+2}{n-2}} - a)}_{II} \; . \tag{1.7.58}$$

But both I and II are strictly negative for $a < 1$, which gives a contradiction, and establishes uniqueness. □

REMARK 1.7.14 The question of stability of solutions of the Lichnerowicz equation has been addressed in [**?**]: surprisingly enough, examples are constructed there where stability fails in dimensions $n \geq 6$.

As a Corollary of Theorem 1.7.13 one obtains:

THEOREM 1.7.15 *Any compact manifold M carries some vacuum initial data set.*

PROOF: We can construct non-trivial solutions of the vector constraint equation using the method of Section 1.7.4, which takes us to the last two columns of the table of Theorem 1.7.13. Choosing some $\tau_\Lambda^2 > 0$ we can then solve the Lichnerowicz equation whatever the Yamabe type of g by the last column in that table. □

As already pointed out, we have the following result, which highlights the importance of Isenberg's Theorem 1.7.6:

PROPOSITION 1.7.16 *All CMC solutions of the vacuum constraint equation can be constructed by the conformal method.*

PROOF: A trivial, albeit not very enlightening proof goes as follows: if (M, g, K) is a CMC vacuum initial data set, the result is established by setting $Y = 0$, $\phi = 1$, $\widetilde{L}^{ij} = K^{ij} - \frac{\mathrm{tr}K}{n} g^{ij}$. □

A natural question is whether the set of solutions to the constraint equations forms a manifold. This was first considered by Fisher and Marsden [67], who provided a Fréchet manifold structure; Banach manifold structures have been constructed in [50], and a Hilbert manifold structure for asymptotically flat initial data sets in [18] (the construction there applies to more general classes of data).

1.7.7 The scalar constraint equation on compact manifolds, $\tau^2 < \frac{2n}{(n-1)}\Lambda$

Theorem 1.7.13 gives an exhaustive description of CMC initial data on compact manifolds when $\tau^2 \geq \frac{2n}{2(n-1)}\Lambda$. Much less is known for Λ's exceeding this bound. As already pointed out, there is observational evidence that Λ might be positive, hence there is direct physical interest for a complete understanding of this case.

When $\tau = 0 = \sigma^2$ but $\Lambda > 0$ we are in the positive case of the Yamabe problem. Obvious examples of three dimensional compact manifolds carrying a metric with positive scalar curvature are given by

$$S^3/\Gamma\,, \quad S^2 \times S^1\,, \tag{1.7.59}$$

where Γ is a discrete subgroup of $O(3)$ without fixed points. The quotient manifolds S^2/Γ are called *spherical manifolds*.

Evolving the time-symmetric vacuum initial data set $(S^n/\Gamma, \mathring{g}_n, 0)$, where $\mathring{g}_n$ denotes a round metric on S^n, one obtains the $(n+1)$-dimensional *de Sitter metrics*:

$$g = -(1 - r^2/\ell^2)dt^2 + \frac{dr^2}{1 - r^2/\ell^2} + r^2 \mathring{g}_{n-1}\,,$$

where $\ell > 0$ is related to the cosmological constant Λ by the formula $2\Lambda = n(n-1)/\ell^2$.

It turns out that a complete description of the possible topologies of three dimensional compact manifolds carrying metrics with positive scalar curvature can be given using the *connected sum* construction, which proceeds as follows: consider any two manifolds M_a, $a = 1, 2$. Consider two sets $B_a \subset M_a$, each diffeomorphic to a ball in $\mathbb{R}^n$. One then defines the manifold $M_1 \# M_2$, called the *connected sum of* M_1 *and* M_2, as the set

$$(M_1 \setminus B_1) \cup ([0,1] \times S^{n-1}) \cup (M_2 \setminus B_2)$$

in which the sphere ∂B_1 is identified in the obvious way with $\{0\} \times S^{n-1}$, and the sphere ∂B_2 is identified with $\{1\} \times S^{n-1}$. In other words, one removes balls from the M_a's and connects the resulting spherical boundaries with a "neck" $[0,1] \times S^{n-1}$.

Consider, then, two manifolds (M_a, g_a) with positive scalar curvature. Gromov and Lawson [80] have shown how to construct a metric of positive scalar curvature on $M_1 \# M_2$. This implies that any compact, orientable three-manifold which is a connected sum of spherical manifolds and of copies of $S^2 \times S^1$ carries a metric of positive scalar curvature. The resolution of the Poincaré conjecture by Perelman [134–136] completes previous work of Schoen-Yau [146] and Gromov-Lawson [81] on this topic, and proves the converse: these are the only compact three-manifolds with positive scalar curvature.

Under the current conditions, a pointwise obstruction to existence of solutions of the Lichnerowicz equation can be derived as follows [83]: Let p_0 be a point where the minimum of ϕ is attained, set $\epsilon := \phi(p_0)$. To emphasize the current sign, we define

$$\Lambda_\tau := \left[-\frac{n-2}{4n}\tau^2 + \frac{n-2}{2(n-1)}\Lambda \right] \geq 0 \; . \tag{1.7.60}$$

At the minimum the Laplacian of ϕ is non-negative, and there the Lichnerowicz equation gives

$$0 \leq \Delta_{\widetilde{g}}\epsilon = \frac{n-2}{4(n-1)}\widetilde{R}\epsilon - \sigma^2 \epsilon^{(2-3n)/(n-2)} - \Lambda_\tau \epsilon^{(n+2)/(n-2)} \tag{1.7.61}$$

But the right-hand-side is negative if $\Lambda_\tau \geq 0$ and if σ^2 is sufficiently large, which gives an obstruction to existence. Setting $a := \epsilon^{4/(n-2)}$, (1.7.61) becomes

$$\begin{aligned} 0 \;&\leq\; \frac{n-2}{4(n-1)}\widetilde{R} - \sigma^2 \epsilon^{-4(n-1)/(n-2)} - \Lambda_\tau \epsilon^{4/(n-2)} \\ &=\; \frac{n-2}{4(n-1)}\widetilde{R} - \sigma^2 a^{-(n-1)} - \Lambda_\tau a \; . \end{aligned}$$

The condition of the vanishing of the derivative with respect to a gives

$$(n-1)\sigma^2 a^{-n} = \Lambda_\tau \; ,$$

and so there *cannot* be a positive minimum of ϕ, and hence a positive solution, if the condition

$$\Lambda_\tau \sigma^{2/(n-1)} \geq (n-1) \left(\frac{n-2}{4n(n-1)} \widetilde{R}_+ \right)^{n/(n-1)} \; , \tag{1.7.62}$$

where $\widetilde{R}_+$ is the positive part of $\widetilde{R}$, holds everywhere. In other words, violation of (1.7.62) *somewhere* is a *necessary* condition for existence of solutions, keeping in mind that equality everywhere would lead to $\epsilon = 0$ as the only possibility for a non-negative minimum of ϕ.

Note that the above calculation can be used to obtain a lower bound on ϕ when σ^2 has no zeros or when $\widetilde{R}$ is strictly positive.

One can also obtain integral, instead of pointwise, conditions for non-existence of solutions, see [83, Theorems 2.1 and 2.2] for details.

In [83, Corollaries 3.1 and 3.2] several criteria for existence have been given, essentially amounting to the requirement that σ^2 be small and without zeros. For example, on compact manifolds such that $\widetilde{R} \geq 0$ but not identically zero, there exists a constant C depending upon $\tilde{g}$ and n such that, if $\Lambda_\tau > 0$ and $\sigma^2 > 0$ and

$$\Lambda_\tau^{n-1} \int_M \sigma^2 \leq C \ ,$$

then a solution exists. This is proved there using the Mountain Pass Lemma [139].

Delaunay metrics

An interesting class of metrics on $S^1 \times S^{n-1}$ with positive scalar curvature is provided by the *Delaunay metrics* which, for $n \geq 3$, take the form

$$g = u^{4/(n-2)}(dy^2 + \mathring{g}_{n-1}) \ , \tag{1.7.63}$$

with $u = u(y)$ and where, as before, $\mathring{g}_{n-1}$ is the unit round metric on S^{n-1}_p. The metrics are spherically symmetric, hence conformally flat. The constant scalar curvature condition $R(g) = n(n-1)$ reduces to an ODE for u:

$$u'' - \frac{(n-2)^2}{4} u + \frac{n(n-2)}{4} u^{\frac{n+2}{n-2}} = 0. \tag{1.7.64}$$

Solutions are determined by two parameters which correspond respectively to a minimum value ε for u, with

$$0 \leq \varepsilon \leq \bar{\varepsilon} = (\frac{n-2}{n})^{\frac{n-2}{4}} \ , \tag{1.7.65}$$

called the *Delaunay parameter* or *neck size*, and a translation parameter along the cylinder. An ODE analysis [121] shows that all the positive solutions are periodic. The degenerate solution with $\varepsilon = 0$ corresponds to the round metric on a sphere from which two antipodal points have been removed. The solution with $\varepsilon = \bar{\varepsilon}$ corresponds to the rescaling of the cylindrical metric so that the scalar curvature has the desired value.

The Delaunay metrics provide an example of countable non-uniqueness of solutions of the Yamabe equation on $S^1 \times S^{n-1}$: for any $T > 0$ and $\ell \in \mathbb{N}^*$ there exists a solution u_ℓ of (1.7.64) with period T/ℓ. Each such function u_ℓ provides a metric with constant scalar curvature $n(n-1)$ on the manifold on which the coordinate y of (1.7.63) is T-periodic.

The ODE (1.7.64) was first studied by Fowler [69, 70], however the name is in analogy with the Delaunay surfaces: the complete, periodic CMC surfaces of revolution in $\mathbb{R}^3$ [61].

Regarding the Delaunay metrics as singular solutions of the Yamabe equation on (S^n_p, g_0) one has a number of uniqueness results. Among these are the facts that no solution with a single singular point exists, and that any solution with exactly two isolated singular points must be conformally equivalent to a Delaunay metric.

It is also known that conformally flat metrics, with constant positive scalar curvature, and with an *isolated singularity of the conformal factor* are necessarily asymptotic to a Delaunay metric [102]; in fact, in dimensions $n = 3, 4, 5$ the conformal flatness condition is not needed [113]. Specifically, in spherical coordinates about an isolated singularity of the conformal factor, there is a half-Delaunay metric which g converges to, exponentially fast in r, along with all of its derivatives. This fact is used in [119–121, 138, 140] where complete, constant scalar curvature metrics, conformal to the round metric on $S_p \setminus \{p_1, \ldots, p_k\}$ were studied and constructed. This is one instance of the more general "singular Yamabe problem".

The time-evolution of time-symmetric Delaunay data leads to the *Kottler–Schwarzschild–de Sitter [103] metrics* in $n+1$ dimensions, with cosmological constant $\Lambda > 0$ and mass $m \in \mathbb{R}$:

$$ds^2 = -V dt^2 + V^{-1} dr^2 + r^2 \mathring{g}_{n-1}, \quad \text{where} \quad V = V(r) = 1 - \frac{2m}{r^{n-2}} - \frac{r^2}{\ell^2} \ , \tag{1.7.66}$$

where $\ell > 0$ is related to the cosmological constant Λ by the formula $2\Lambda = n(n-1)/\ell^2$. Comparing (1.7.66) and (1.7.63) we find

$$r = u^{\frac{2}{n-2}} \ , \quad r\frac{dy}{dr} = V^{-1/2} \ , \tag{1.7.67}$$

which allows us to determine y as a function of r on any interval of r's on which V has no zeros.

To avoid a singularity lying at finite distance on the level sets of t one needs $m > 0$. Equation (1.7.66) provides then a spacetime metric satisfying the Einstein equations with cosmological constant $\Lambda > 0$ and with well behaved spacelike hypersurfaces when one restricts the coordinate r to an interval (r_b, r_c) on which $V(r)$ is positive; such an interval exists if and only if

$$\left(\frac{2}{(n-1)(n-2)}\right)^{n-2} \Lambda^{n-2} m^2 n^2 < 1 \ . \tag{1.7.68}$$

When $n = 3$ this corresponds to the condition that $9m^2\Lambda < 1$.

1.7.8 Matter fields

The conformal method easily extends to CMC constraint equations for some non-vacuum initial data, e.g. the Einstein-Maxwell system [87] where one obtains results very similar to those of Theorem 1.7.13. However, other important examples, such as the Einstein-scalar field system [39–41, 83], require more effort and are not as fully understood.

Recall that the energy density μ and the energy-momentum density J of matter fields is related to the geometry through the formulae

$$16\pi\mu \ := \ R(g) - |K|^2_g + (\mathrm{tr}_g K)^2 - 2\Lambda \tag{1.7.69}$$

$$8\pi J^i \ := \ D_i(K^{ij} - \mathrm{tr}_g K g^{ij}) \ . \tag{1.7.70}$$

If μ and J have been prescribed, this becomes a system of equations for ϕ and $\widetilde{L}$

$$\begin{aligned}\Delta_{\widetilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi &= -\sigma^2\phi^{\frac{2-3n}{n-2}} + \beta\phi^{\frac{n+2}{n-2}} \\ &\quad -\frac{4(n-2)}{(n-1)}\phi^{\frac{n+2}{n-2}}\pi\mu \,. \qquad (1.7.71)\\ \tilde{D}_i\widetilde{L}^{ij} &= 8\pi\phi^{\frac{2(n+2)}{n-2}}J^i + \frac{n-1}{n}\phi^{\frac{2n}{n-2}}\tilde{D}^j\tau \,. \qquad (1.7.72)\end{aligned}$$

It is important to realize that the conformal method has no physical contents, and is an ansatz for constructing solutions of the constraint equations. The question of scaling properties of μ and J under conformal transformations is thus largely a matter of convenience. For instance, *when* $d\tau = 0$, a convenient prescription for J is to set

$$\tilde{J}^i = \phi^{\frac{2(n+2)}{n-2}}J^i \,, \qquad (1.7.73)$$

and to view $\tilde{J}$ as free data, for then (1.7.72) decouples from (1.7.71). There is then a natural rescaling of μ which arises from the *dominant energy condition* $\mu^2 \geq g_{ij}J^iJ^j$: since $g_{ij} = \phi^{\frac{4}{n-2}}\widetilde{g}_{ij}$, under (1.7.73) we have

$$g_{ij}J^iJ^j = \phi^{\frac{4}{n-2}-\frac{4(n+2)}{n-2}}\widetilde{g}_{ij}\tilde{J}^i\tilde{J}^j = \phi^{-\frac{4(n+1)}{n-2}}\widetilde{g}_{ij}\tilde{J}^i\tilde{J}^j \,,$$

and so the dominant energy condition will be covariant under these rescalings if we set

$$\tilde{\mu} = \phi^{\frac{2(n+1)}{n-2}}\mu \,, \qquad (1.7.74)$$

viewing $\tilde{\mu}$ as the free data, and μ as the derived ones. The scaling (1.7.73)-(1.7.74) is known to us from [35], where it has been termed *York scaling*. With those definitions (1.7.71)-(1.7.72) become

$$\begin{aligned}\Delta_{\widetilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi &= -\sigma^2\phi^{\frac{2-3n}{n-2}} + \beta\phi^{\frac{n+2}{n-2}} \\ &\quad -\frac{4(n-2)}{(n-1)}\phi^{-\frac{n}{n-2}}\pi\tilde{\mu} \,. \qquad (1.7.75)\\ \tilde{D}_i\widetilde{L}^{ij} &= 8\pi\tilde{J}^i + \frac{n-1}{n}\phi^{\frac{2n}{n-2}}\tilde{D}^j\tau \,. \qquad (1.7.76)\end{aligned}$$

1.7.9 Maxwell fields

Let the (spacelike) initial data hypersurface $\mathscr{S}$ be given by the equation $x^0 = 0$, define

$$\nu := \frac{1}{\sqrt{-g^{00}}} \,, \qquad (1.7.77)$$

so that the future directed unit normal N to $\mathscr{S}$ has covariant components

$$N_\mu dx^\mu = -\nu dx^0 \,.$$

When constructing initial data involving Maxwell equations, one needs to keep in mind that the Maxwell equations

$$\nabla_\mu F^{\mu\nu} = 4\pi J^\nu_M \,, \qquad \nabla_\mu \star F^{\mu\nu} = 0 \,, \qquad (1.7.78)$$

where $\star F^{\mu\nu} := \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$, and where J^μ_M denotes the charge-current four-vector density, imply constraints equations on the electric and magnetic fields E^i and B^i (compare (1.6.2), p. 23):

$$E^i\partial_i := N_\mu F^{i\mu}\partial_i = \nu F^{0i}\partial_i \;, \qquad B^i\partial_i := \nu \star F^{i0}\partial_i \;. \tag{1.7.79}$$

Indeed, we then have

$$D_iE^i = 4\pi\nu J^0_M \;, \qquad D_iB^i = 0 \;. \tag{1.7.80}$$

Since $D_iB^i = \partial_i(\sqrt{\det g_{k\ell}}B^i)/\sqrt{\det g_{k\ell}}$, the second equation above is immediately covariant under conformal rescalings of the metric if B^i is taken of the form

$$B^i := \phi^{-\frac{2n}{n-2}}\tilde B^i \;, \tag{1.7.81}$$

where $\tilde B^i$ is divergence-free in the metric $\tilde g$. Likewise the first equation in (1.7.80) will be conformally covariant if

$$E^i := \phi^{-\frac{2n}{n-2}}\tilde E^i \;, \quad \nu J^i_M := \phi^{-\frac{2n}{n-2}}\tilde\nu\tilde J^i_M \;, \quad \tilde D_i\tilde E^i = 4\pi\tilde\nu\tilde J^0_M \;. \tag{1.7.82}$$

The energy-density of the Maxwell fields is (keeping in mind that $n = 3$ here)

$$\begin{aligned}\mu &:= \frac{1}{8\pi}\left(g_{ij}E^iE^j + g_{ij}B^iB^j\right)\\ &= \varphi^{-8}\underbrace{\frac{1}{8\pi}\left(\widetilde g_{ij}\tilde E^i\tilde E^j + \widetilde g_{ij}\tilde B^i\tilde B^j\right)}_{=:\tilde\mu}\;.\end{aligned} \tag{1.7.83}$$

One recovers the York scaling (1.7.74). One similarly checks the York-scaling property for the energy-momentum three-vector J^i.

1.8 Non-compact initial data sets: an overview

So far we have been considering initial data sets on compact manifolds. However, there are *noncompact* classes of data which are of interest, and in this section we will review them.

Recall that a vacuum initial data set (M, g, K) is a triple consisting of an n-dimensional manifold M, a Riemannian metric on g, and a symmetric two-covariant tensor K. One moreover requires that the vacuum constraint equations hold:

$$R(g) = 2\Lambda + |K|^2_g - (\mathrm{tr}_g K)^2 \;, \tag{1.8.1}$$

$$D_jK^j{}_k - D_kK^j{}_j = 0 \;. \tag{1.8.2}$$

The reader will note that we have allowed for a non-zero cosmological constant $\Lambda \in \mathbb{R}$. The hypothesis that $\Lambda = 0$ is adequate when describing isolated gravitating systems such as the solar system; $\Lambda > 0$ seems to be needed in cosmology in view of the observations of the rate of change of the Hubble constant [142, 155]; finally, a negative cosmological constant appears naturally in many models of theoretical physics, such as string theory or supergravity. For those reasons it is of interest to consider all possible values of Λ.

A *CMC initial data set* is one where $\mathrm{tr}_g K$ is constant; data are called *maximal* if $\mathrm{tr}_g K$ is identically zero. A *time-symmetric* data set if one where K vanishes identically. In this case (1.8.1)-(1.8.2) reduces to the requirement that the scalar curvature of g be constant:

$$R(g) = 2\Lambda \ .$$

1.8.1 Non-compact manifolds with constant positive scalar curvature

The topological classification of compact three-manifolds with positive scalar curvature generalises to the following non-compact setting: One says that a Riemannian metric g has bounded geometry if g has bounded sectional curvatures and injectivity radius bounded away from zero. Using Ricci flow (the short-time existence of which is guaranteed in the setting by the work of Shi [149]), one has [26]:

THEOREM 1.8.1 *Let $\mathscr{S}$ be a connected, orientable three-manifold which carries a complete Riemannian metric of bounded geometry and uniformly positive scalar curvature. Then there is a finite collection $\mathcal{F}$ of spherical manifolds such that $\mathscr{S}$ is an (infinite) connected sum of copies of $S^1 \times S^2$ and members of $\mathcal{F}$.*

The cylinders $(\mathbb{R}\times S^{n-1}, dx^2+\mathring{g}_{n-1})$ provide examples of non-compact manifolds with positive scalar curvature. The underlying manifold can be viewed as S^n from which the north and south poles have been removed. In Theorem 1.8.1 they are viewed as an infinite connected sum of $S^1 \times S^2$.

Completing the initial data $(\mathbb{R} \times S^2, dx^2 + \mathring{g}_2)$ with a suitable extrinsic curvature tensor, when $\Lambda = 0$ the corresponding evolution leads to the interior Schwarzschild metric: for $t < 2m$

$$g = -\frac{1}{\frac{2m}{t} - 1}dt^2 + \left(\frac{2m}{t} - 1\right) dx^2 + t^2 \mathring{g}_2 \ . \qquad (1.8.3)$$

When $\Lambda > 0$, vanishing extrinsic curvature leads to the *Nariai metrics* [128]:

$$g = \frac{1}{\Lambda}\left(-dt^2 + \cosh^2 t \, d\rho^2 + \mathring{g}_2\right) \qquad (1.8.4)$$

(compare [29]).

Another class of positive scalar constant curvature metrics on $S^1 \times S^{n-1}$ is provided by the *Delaunay metrics*, when the coordinate y of (1.7.63) is not periodically identified, but runs over $\mathbb{R}$.

EXAMPLE 1.8.2 The *Delaunay* metrics provide an example of complete spherically symmetric metrics with positive scalar curvature. Large classes of metrics with the last set of properties can be constructed as follows: Recall that (see (B.1.14), p. 112, with g and $\tilde{g}$ there interchanged)

$$g_{ij} = \varphi^{\frac{4}{n-2}}\tilde{g}_{ij} \quad \Longrightarrow \quad R = \varphi^{-\frac{4}{n-2}}\left(\widetilde{R} - \frac{4(n-1)}{(n-2)\varphi}\Delta_{\tilde{g}}\varphi\right) \ . \qquad (1.8.5)$$

So if $\tilde{g}$ is the flat Euclidean metric δ, then g will have non-negative scalar curvature if and only if

$$\Delta_\delta \varphi \le 0 \ .$$

Smooth pherically symmetric solutions of this inequality which are regular at the origin and which asymptote to one at infinity can be obtained by setting

$$\varphi = 1 + \frac{1}{r^{n-2}} \int_0^r f(s) s^{n-1} ds + \int_r^\infty f(t) t dt \ , \tag{1.8.6}$$

where f is any smooth positive function such that $\int_0^\infty f(r) r^{n-1} dr$ is finite. (The solution will be be asymptotically flat in the usual sense if, e.g., f is compactly supported.) Indeed, we have

$$\varphi' = -\frac{(n-2)}{r^{n-1}} \int_0^r f(s) s^{n-1} ds \ , \tag{1.8.7}$$

hence

$$\Delta_\delta \varphi = r^{-(n-1)} \partial_r (r^{n-1} \partial_r \varphi) = -(n-2) f \ ,$$

and so the sign of the scalar curvature of $\varphi^{4/(n-2)}\delta$ is determined by that of $-f$.

In the region where f vanishes we have $R = 0$, which provides vacuum regions. Connected regions of non-zero f can be thought of as a central star, or shells of matter.

It is interesting to enquire about existence of spherically symmetric *minimal surfaces* for such metrics. Now a strict definition of a minimal surface is the requirement of minimum area amongst nearby competing surfaces, but any critical point of the area functional is also often called "minimal", and we will follow this practice.

In the case under consideration, the area of spheres of constant radius is proportional to $(r^2 \varphi^{4/(n-2)})^{(n-1)/2}$, and so the area will have vanishing derivative with respect to r if and only if

$$\begin{aligned} (r\varphi^{2/(n-2)})' = 0 \quad &\Longleftrightarrow \quad -\frac{2r\varphi'}{(n-2)} = \varphi \\ &\Longleftrightarrow \quad \underbrace{\frac{1}{r^{n-2}} \int_0^r f(s) s^{n-1} ds}_{=:h(r)} = \underbrace{1 + \int_r^\infty f(t) t dt}_{=:g(r)} \ . \end{aligned} \tag{1.8.8}$$

This formula can be used to construct solutions containing minimal spheres as follows: Suppose that the function f of (1.8.6) is constant, say $f = f_0 > 0$, on an interval $[0, r_1]$. Then $h(r) = f_0 r^2/n$ increases while $g(r) = g(0) - f_0 r^2/2$ decreases, so equality will be achieved precisely once somewhere before r_1 if

$$\frac{f_0 r_1^2}{n} > 1 + \int_{r_1}^\infty f(r) r dr \ . \tag{1.8.9}$$

The value of $r \in [0, r_1]$ at which the equality in (1.8.8) holds provides our first "minimal" surface (which actually locally maximizes area). We then let f drop smoothly to zero on $[r_1, r_2]$, for some $r_2 > r_1$, and keep f equal to zero on $[r_2, r_3)$, with r_3 possibly equal to ∞. On $[r_1, r_2]$ the function h continues to increase while the function g continues to decrease, so there cannot be any further minimal surfaces in this interval. On $[r_2, r_3)$ the function

$$h(r) = r^{-n-2} \int_0^{r_2} f(s) s^{n-1} ds$$

decreases as $r^{-(n-2)}$ while $g(r)$ remains constant, so equality in (1.8.8) will be attained before r_3 if

$$r_3^{-n-2} \int_0^{r_2} f(s) s^{n-1} ds < 1 + \int_{r_3}^{\infty} f(r) r \, dr \ . \tag{1.8.10}$$

Note that the choice of r_3 does not affect (1.8.9) insofar as f vanishes on $[r_2, r_3)$.

In particular if we choose $r_3 = \infty$, we can first choose any central value f_0, and then choose r_1 so that $f_0 r_1^2/n > 2$. Choosing the intermediate region $[r_1, r_2]$ small enough so that the integral at the right-hand-side is smaller than one, it follows from (1.8.9) that the resulting metric will have precise one locally maximal sphere somewhere before r_1. For $r > r_2$ the metric is the (asymptotically flat) space-Schwarzschild metric with a second minimal sphere somewhere in $[r_2, \infty)$.

1.8.2 Asymptotically flat manifolds

One of the most widely studied class of Lorentzian manifolds are the *asymptotically flat space-times* which model isolated gravitational systems. Now, there exist several ways of defining asymptotic flatness, all of them roughly equivalent in vacuum. In this section we describe the Cauchy data point of view, which appears to be the least restrictive in any case.

So, a space-time $(\mathscr{M}, \mathfrak{g})$ will be said to possess an *asymptotically flat end* if $\mathscr{M}$ contains a spacelike hypersurface $M_{\rm ext}$ diffeomorphic to $\mathbb{R}^n \setminus B(R)$, where $B(R)$ is a coordinate ball of radius R. An end comes thus equipped with a set of Euclidean coordinates $\{x^i, i = 1, \ldots, n\}$, and one sets $r = |x| := \big(\sum_{i=1}^n (x^i)^2\big)^{1/2}$. One then assumes that there exists a constant $\alpha > 0$ such that, in local coordinates on $M_{\rm ext}$ obtained from $\mathbb{R}^n \setminus B(R)$, the metric g induced by $\mathfrak{g}$ on $M_{\rm ext}$, and the second fundamental form K of $M_{\rm ext}$, satisfy the fall-off conditions, for some $k > 1$,

$$g_{ij} - \delta_{ij} = O_k(r^{-\alpha}) \ , \qquad K_{ij} = O_{k-1}(r^{-1-\alpha}) \ , \tag{1.8.11}$$

where we write $f = O_k(r^\beta)$ if f satisfies

$$\partial_{k_1} \ldots \partial_{k_\ell} f = O(r^{\beta - \ell}) \ , \quad 0 \le \ell \le k \ . \tag{1.8.12}$$

The PDE aspects of the problem require furthermore (g, K) to lie in certain weighted Hölder or Sobolev spaces defined on $\mathscr{S}$. More precisely, the above decay conditions should be implemented by conditions on the Hölder continuity of the fields; alternatively, the above equations should be understood in an integral sense. The constraint equations can be conveniently treated in both Hölder and Sobolev spaces, but one should keep in mind that L^2-type Sobolev spaces are better suited for solving the evolution equations.

The analysis of elliptic operators such as the Laplacian on weighted functional spaces was initiated by Nirenberg and Walker [129]; see also [13, 36, 109–112, 122–124, 131] as well as [35].

The conformal method works again very well for asymptotically flat initial data sets. The approach is very similar to the one for compact manifolds, with two important distinctions: on non-compact manifolds the embeddings

$H_k \subset H_m$ and $C^{k,\alpha} \subset C^{m,\alpha}$, for $k \geq m$, are not compact anymore. Furthermore, to obtain good mapping properties for elliptic operators one needs to introduce weighted Sobolev or Hölder spaces. The reader is referred to the original references for details [33, 35, 36, 42, 43, 45, 116, 117].

CMC initial data can only be asymptotically flat if $\Lambda = \tau = 0$. The Lichnerowicz equation simplifies then to

$$\Delta_{\tilde{g}}\phi - \frac{n-2}{4(n-1)}\widetilde{R}\phi = -\sigma^2\phi^{(2-3n)/(n-2)} , \tag{1.8.13}$$

where

$$\sigma^2 := \frac{n-2}{4(n-1)}|\widetilde{L}|^2_{\tilde{g}} . \tag{1.8.14}$$

The treatment of TT-tensors is essentially identical to that on compact manifolds. In fact, the analysis is somewhat simpler because there are no conformal Killing vectors which decay to zero as one recedes to infinity, so the conformal vector Laplacian has no kernel on weighted Sobolev spaces with decay.

Concerning the Lichnerowicz equation (1.8.13), suppose that there exists a positive solution of this equation, then the conformally rescaled metric $g = \phi^{4/(n-2)}\tilde{g}$ has non-negative scalar curvature $R = |L|^2_g$, with L being an appropriate rescaling of $\widetilde{L}$. Thus, a necessary condition for existence of positive solutions of (1.8.13) is that there exist metrics of non-negative scalar curvature in the conformal class $[\tilde{g}]$ of $\tilde{g}$. A precise conformally invariant criterion for this has been proposed in [33] but, as emphasised e.g. in [73, 116], the statement in [33] is not quite correct. In [73, 116] a corrected version has been provided, as follows:

Recall that the *Yamabe number* of a metric is defined by the equation

$$Y(M,g) = \inf_{u\in C^\infty_b,\ u\not\equiv 0} \frac{\int_M(|Du|^2 + \frac{n-2}{4(n-1)}Ru^2)}{\left(\int_M u^{2n/(n-2)}\right)^{(n-2)/n}} . \tag{1.8.15}$$

where C^∞_b denotes the space of compactly supported smooth functions. As discussed in Section 1.7.1, $Y(M,g)$ depends only upon the conformal class of g. For asymptotically flat manifolds, there exists a conformal rescaling so that $\widetilde{R}$ is non-negative if and only if $Y(M,g) > 0$ [33, 73, 116].

Suppose, then, that we can perform the conformal rescaling that makes $\widetilde{R} \geq 0$. Setting

$$\phi = 1 + u ,$$

the requirement that g has vanishing scalar curvature translates into an equation for u:

$$\Delta_{\tilde{g}}u - \frac{n-2}{4(n-1)}\widetilde{R}u = -\frac{n-2}{4(n-1)}\widetilde{R} . \tag{1.8.16}$$

Because $\widetilde{R}$ is non-negative now, there is no difficulty in finding a solution u decaying to zero at infinity, with suitable weighted regularity. Note that u is strictly positive by the maximum principle, in particular $1 + u$ has no zeros and asymptotes to one. Replacing $\tilde{g}$ by $(1+u)^{4/(n-2)}\tilde{g}$, the new $\tilde{g}$ is again asymptotically flat, and (1.8.13) simplifies to

$$\Delta_{\tilde{g}}\phi = -\sigma^2\phi^{(2-3n)/(n-2)} . \tag{1.8.17}$$

Since

$$0 = \Delta_{\tilde{g}} 1 \geq -\sigma^2 = -\sigma^2 1^{(2-3n)/(n-2)} ,$$

the constant function

$$\varphi_- := 1$$

provides a subsolution which asymptotes to one. To obtain a supersolution we use a variation of Hebey's trick, as used in our treatment of compact manifolds: Let v be a solution approaching zero in the asymptotically flat regions of

$$\Delta_{\tilde{g}} v = -\sigma^2 . \tag{1.8.18}$$

Then v is strictly positive by the maximum principle. Let $\phi_+ = 1 + v$, then

$$\Delta_{\tilde{g}} \phi_+ = -\sigma^2 \leq -\sigma^2 (1+v)^{(2-3n)/(n-2)} = -\sigma^2 \phi_+^{(2-3n)/(n-2)} , \tag{1.8.19}$$

so ϕ_+ is indeed a supersolution. Since $\phi_+ := 1 + v \geq 1 =: \phi_-$, we can use the monotone iteration scheme to obtain a solution. Note that ϕ_- and ϕ_+ both asymptote to one, so the solution also will.

This provides a complete description of vacuum, asymptotically flat, CMC initial data.

1.9 Non-CMC data

One can also consider the conformal method *without* assuming CMC data. As before, the free conformal data consist of a manifold M, a Riemannian metric $\tilde{g}$ on M, a trace-free symmetric tensor $\tilde{\sigma}$, and a *mean curvature function* τ. The fields (g, K) defined as

$$g = \phi^q \tilde{g} , \qquad \text{where } q = \frac{4}{n-2} , \tag{1.9.1}$$

$$K = \phi^{-2}(\tilde{\sigma} + \tilde{C}(Y)) + \frac{\tau}{n} \phi^q g , \tag{1.9.2}$$

where ϕ is positive, will then solve the constraint equations with matter energy-momentum density (μ, J) if and only if the function ϕ and the vector field Y solve the equations

$$\mathrm{div}_{\tilde{g}}(\tilde{C}(Y) + \tilde{\sigma}) = \frac{n-1}{n} \phi^{q+2} \widetilde{D}\tau + 8\pi \phi^{q_J} \tilde{J} , \tag{1.9.3}$$

$$\Delta_{\tilde{g}} \phi - \frac{1}{q(n-1)} R(\tilde{g}) \phi + \frac{1}{q(n-1)} |\tilde{\sigma} + \tilde{C}(Y)|^2_{\tilde{g}} \phi^{-q-3} - \frac{1}{qn} \tau^2 \phi^{q+1} = 16\pi \phi^{q_\mu} \tilde{\mu} . \tag{1.9.4}$$

Here q_J and q_μ are exponents which can be chosen in a manner which is convenient for the problem at hand. A possible choice is obained by inserting the *York scaling* given in (1.7.73)-(1.7.74) into (1.7.71)-(1.7.72); this is convenient e.g. for the Einstein-Maxwell constraints in dimension $n = 3$. Finally, the symbol $\widetilde{D}$ denotes the covariant derivative of $\tilde{g}$, and $\tilde{C}(Y)$ is the *conformal Killing operator* of $\tilde{g}$:

$$\tilde{C}(Y)_{ab} = \widetilde{D}_a Y_b + \widetilde{D}_b Y_a - \frac{2}{n} \tilde{g}_{ab} \widetilde{D}_c Y^c . \tag{1.9.5}$$

When $d\tau \not\equiv 0$, the vector constraint equation does not decouple from the scalar one, and one needs to find simultaneously the solution (ϕ, Y) to both equations above.

One can invoke the implicit function theorem to construct solutions of the above when τ is bounded away from zero and $d\tau$ is sufficiently small, near a solution at which the linearized operator is an isomorphism. Other techniques have also been used in this context in [3, 38, 93, 94]. A non-existence theorem for a class of near-CMC conformal data has been established in [95].

The first general result without assuming small gradient is due to Holst, Nagy, and Tsogtgerel [85, 86] who assumed non-vanishing matter source, $\tilde\mu \not\equiv 0$. Maxwell [118] has extended their argument to include the vacuum case, leading to:

THEOREM 1.9.1 (Holst, Nagy, Tsogtgerel [85, 86], Maxwell [118]) *Let $(M, \widetilde{g}_{ab})$ be a three dimensional, smooth, compact Riemannian manifold of positive Yamabe type without conformal Killing vectors, and let $\tilde\sigma^{ij}$ be a symmetric transverse traceless tensorfield. If the seed tensor $\tilde\sigma$ and the matter sources $|\tilde J|_{\widetilde g} \le \tilde\mu$ are sufficiently small, then there exists a scalar field $\phi > 0$ and a vector field Y solving the system*

$$\begin{aligned}\Delta_{\widetilde g}\phi - \tfrac{1}{8}\widetilde R\phi &= -\tfrac{1}{8}|\tilde\sigma|^2_{\widetilde g}\phi^{-7} + \tfrac{1}{12}\tau^2\phi^5 - 2\pi\tilde\mu\phi^{-3}\ ,\\ \tilde D_i(\widetilde D^i Y^j + \widetilde D^j Y^i - \tfrac{2}{3}\widetilde D_k Y^k \widetilde g^{ij}) &= 8\pi \tilde J^j + \tfrac{2}{3}\phi^6 \tilde D^j\tau\ ,\end{aligned} \tag{1.9.6}$$

and hence providing a solution

$$(g_{ij}, K^{ij}) = (\phi^4\widetilde g_{ij}, \phi^{-10}(\tilde\sigma^{ij} + \widetilde D^i Y^j + \widetilde D^j Y^i - \frac{2}{3}\widetilde D_k Y^k \widetilde g^{ij}) + \frac{\tau}{3}\phi^{-4}\widetilde g^{ij})$$

of the constraint equations in vacuum ($\mu = 0 = J$) or with sources

$$(\mu, J^i) = (\phi^{-\frac{2(n+1)}{n-2}}\tilde\mu, \phi^{-\frac{2(n+2)}{n-2}}\tilde J^i)$$

(compare (1.7.74) and (1.7.73), p. 48).

The reader is referred to [118] for further general statements concerning the problem at hand.

1.10 Gluing techniques

The *gluing techniques* can be regarded as a singular perturbation method. They are used to produce new solutions by gluing together old ones. The main usefulness of the technique in general relativity lies in the fact that, away from the small set about which one fuses the two solutions, the new solution is identical, to the original ones. This gives one control on the physical properties of the glued initial data. Furthermore, because of the finite speed of propagation of signals, one has information of the global behaviour in time of the resulting solutions, at least in some regions.

The linearized constraint equations and KIDs

The starting point of gluing constructions for the constraint equations is the linearization of these equations about a given solution (M, g, K). We let $\mathcal{P}^*_{(g,K)}$ denote the L^2 adjoint of the linearization of the constraint equations at this solution. Viewed as an operator acting on a scalar function N and a vector field Y, $\mathcal{P}^*_{(g,K)}$ takes the explicit form [49]

$$\mathcal{P}^*_{(g,K)}(N,Y) = \left(\begin{array}{l} 2(D_{(i}Y_{j)} - D^lY_lg_{ij} - K_{ij}N + \mathrm{tr}K\, Ng_{ij}) \\ \\ D^lY_lK_{ij} - 2K^l{}_{(i}D_{j)}Y_l + K^q{}_lD_qY^lg_{ij} \\ \ -\Delta Ng_{ij} + D_iD_jN + (D^pK_{lp}g_{ij} - D_lK_{ij})Y^l \\ \ -N\mathrm{Ric}\,(g)_{ij} + 2NK^l{}_iK_{jl} - 2N(\mathrm{tr}\,K)K_{ij} \end{array} \right) . \quad (1.10.1)$$

Now this operator does not, on first inspection, appear to be very "user friendly". However, our immediate concern is solely with its kernel, and the pairs (N, Y) which lie in its kernel have a very straightforward geometric and physical characterization. In particular, let Ω be an open subset of M. By definition, the set of "KIDs" on Ω, denoted $\mathscr{K}(\Omega)$, is the set of all solutions of the equation

$$\mathcal{P}^*_{(\gamma,K)|_\Omega}(N,Y) = 0\ . \quad (1.10.2)$$

Such a solution (N, Y), if nontrivial, generates a space-time Killing vector field in the domain of dependence of $(\Omega, g|_\Omega, K|_\Omega)$ [126].

From a geometric point of view one expects that solutions with symmetries should be rare. This was made rigorous in [24], where it is shown that the generic behaviour among solutions of the constraint equations is the absence of KIDs on any open set. On the other hand, one should note that essentially every explicit solution has symmetries. In particular, both the flat initial data for Minkowski space, and the initial data representing the constant time slices of Schwarzschild have KIDs.

Corvino's result

As we have already pointed out, the Einstein constraint equations form an underdetermined system of equations, and as such, it is unreasonable to expect that they (or their linearizations) should satisfy the unique continuation property. In 2000, Corvino established a gluing result for asymptotically flat metrics with zero scalar curvature which dramatically illustrated this point [56]. In the special case when one considers initial data with vanishing second fundamental form $K \equiv 0$, the momentum constraint equation becomes trivial and the Hamiltonian constraint equation reduces to simply $R(g) = 0$, i.e. a scalar flat metric. Such initial data sets are referred to as "time-symmetric" because the space-time obtained by evolving them possesses a time-reversing isometry which leaves the initial data surface fixed. Beyond Euclidean space itself, the constant time slices of the Schwarzschild space-time form the most basic examples of asymptotically flat, scalar flat manifolds. One long-standing open problem [16, 147] in the field had been whether there exist scalar flat metrics on $\mathbb{R}^n$ which are not globally spherically symmetric but which are spherically symmetric in a neighborhood of infinity and hence, by Birkhoff's theorem, Schwarzschild there.

Corvino resolved this by showing that he could deform any asymptotically flat, scalar flat metric to one which is exactly Schwarzschild outside of a compact set.

Theorem 1.10.1 ([56]) *Let (M, g) be a smooth Riemannian manifold with zero scalar curvature containing an asymptotically flat end $\mathscr{S}_{\mathrm{ext}} = \{|x| > r > 0\}$. Then there is a $R > r$ and a smooth metric $\bar{g}$ on M with zero scalar curvature such that $\bar{g}$ is equal to g in $M \setminus \mathscr{S}_{\mathrm{ext}}$ and $\bar{g}$ coincides on $\{|x| > R\}$ with the metric induced on a standard time-symmetric slice in the Schwarzschild solution. Moreover the mass of $\bar{g}$ can be made arbitrarily close to that of g by choosing R sufficiently large.*

Underlying this result is a gluing construction where the deformation has compact support. The ability to do this is a reflection of the underdetermined nature of the constraint equations. In this setting, since $K \equiv 0$, the operator takes a much simpler form, as a two-covariant tensor valued operator acting on a scalar function u by

$$\mathcal{P}^* u = -(\Delta_g u) g + \mathrm{Hess}_g u - u \mathrm{Ric}(g) \ .$$

An elementary illustration of how an underdetermined system can lead to compactly supported solutions is given by the construction of compactly supported transverse-traceless tensors on $\mathbb{R}^3$ in Appendix B of [57] (see also [22, 60]).

An additional challenge in proving Theorem 1.10.1 is the presence of KIDs on the standard slice of the Schwarzschild solution. If the original metric had ADM mass $m(g)$, a naive guess could be that the best fitting Schwarzschild solution would be the one with precisely the same mass. However the mass, and the coordinates of the center of mass, are in one-to-one correspondence with obstructions arising from KIDs. To compensate for this co-kernel in the linearized problem, Corvino uses these ($n+1$ in dimension n) degrees of freedom as effective parameters in the geometric construction. The final solution can be chosen to have its ADM mass arbitrarily close to the initial one.

Corvino's technique has been applied and extended in a number of important ways. The "asymptotic simplicity" model for isolated gravitational systems proposed by Penrose [133] has been very influential. This model assumes existence of smooth conformal completions to study global properties of asymptotically flat space-times. The question of existence of such vacuum space-times was open until Chruściel and Delay [48], and subsequently Corvino [57], used this type of gluing construction to demonstrate the existence of infinite dimensional families of vacuum initial data sets which evolve to asymptotically simple space-times. The extension of the gluing method to non-time-symmetric data was done in [49, 58]. This allowed for the construction of space-times which are exactly Kerr outside of a compact set, as well as showing that one can specify other types of useful asymptotic behavior.

Conformal gluing

In [91], Isenberg, Mazzeo and Pollack developed a gluing construction for initial data sets satisfying certain natural non-degeneracy assumptions. The perspective taken there was to work within the conformal method, and thereby establish a gluing theorem for solutions of the determined system of PDEs given by (1.9.3) and (1.9.4). This was initially done only within the setting of constant mean curvature initial data sets and in dimension $n = 3$ (the method was extended to all higher dimensions in [89]). The construction of [91] allowed one to combine initial data sets by taking a connected sum of their underlying manifolds, to add wormholes (by performing codimension 3 surgery on the underlying, connected, 3-manifold) to a given initial data set, and to replace arbitrary small neighborhoods of points in an initial data set with asymptotically hyperbolic ends.

In [92] this gluing construction was extended to only require that the mean curvature be constant in a small neighborhood of the point about which one wanted to perform a connected sum. This enabled the authors to show that one can replace an arbitrarily small neighborhood of a generic point in any initial data set with an asymptotically *flat* end. As we have seen that CMC solutions of the vacuum constraint equations exist on any compact manifold, this leads to the following result, which asserts that there are no topological obstructions to asymptotically flat solutions of the constraint equations:

Theorem 1.10.2 ([92]) *Let M be any closed n-dimensional manifold, and $p \in M$. Then $M \setminus \{p\}$ admits an asymptotically flat initial data set satisfying the vacuum constraint equations.*

Initial data engineering

The gluing constructions of [91] and [92] are performed using a determined elliptic system provided by the conformal method, which necessarily leads to a global deformation of the initial data set, small away from the gluing site. Now, the ability of the Corvino gluing technique to establish compactly supported deformations invited the question of whether these conformal gluings could be localized. This was answered in the affirmative in [49] for CMC initial data under the additional, generically satisfied [24], assumption that there are no KIDs in a neighborhood of the gluing site.

In [52, 53], this was substantially improved upon by combining the gluing construction of [91] together with the Corvino gluing technique of [48, 56], to obtain a localized gluing construction in which the only assumption is the absence of KIDs near points. For a given n-manifold M (which may or may not be connected) and two points $p_a \in M$, $a = 1, 2$, we let $\tilde{M}$ denote the manifold obtained by replacing small geodesic balls around these points by a neck $S^{n-1} \times I$. When M is connected this corresponds to performing codimension n surgery on the manifold. When the points p_a lie in different connected components of M, this corresponds to taking the connected sum of those components.

THEOREM 1.10.1 ([52, 53]) *Let (M, g, K) be a smooth vacuum initial data set, with M not necessarily connected, and consider two open sets $\Omega_a \subset M$, $a = 1, 2,$*

with compact closure and smooth boundary such that

$$\text{the set of KIDs, } \mathscr{K}(\Omega_a)\text{, is trivial.}$$

Then for all $p_a \in \Omega_a$, $\epsilon > 0$, *and* $k \in \mathbb{N}$ *there exists a smooth vacuum initial data set* $(\tilde{M}, g(\epsilon), K(\epsilon))$ *on the glued manifold* $\tilde{M}$ *such that* $(g(\epsilon), K(\epsilon))$ *is* ϵ*-close to* (g, K) *in a* $C^k \times C^k$ *topology away from* $B(p_1, \epsilon) \cup B(p_2, \epsilon)$*. Moreover* $(g(\epsilon), K(\epsilon))$ *coincides with* (g, K) *away from* $\Omega_1 \cup \Omega_2$*.*

This result is sharp in the following sense: first note that, by the positive mass theorem, initial data for Minkowski space-time cannot locally be glued to anything which is non-singular and vacuum. This meshes with the fact that for Minkowskian initial data, we have $\mathscr{K}(\Omega) \neq \{0\}$ for any open set Ω. Next, recall that by the results in [24], the no-KID hypothesis in Theorem 1.10.1 is generically satisfied. Thus, the result can be interpreted as the statement that for generic vacuum initial data sets the local gluing can be performed around arbitrarily chosen points p_a. In particular the collection of initial data with generic regions Ω_a satisfying the hypotheses of Theorem 1.10.1 is not empty.

The proof of Theorem 1.10.1 is a mixture of gluing techniques developed in [89, 91] and those of [49, 56, 58]. In fact, the proof proceeds initially via a generalization of the analysis in [91] to compact manifolds with boundary. In order to have CMC initial data near the gluing points, which the analysis based on [91] requires, one makes use of the work of Bartnik [14] on the plateau problem for prescribed mean curvature spacelike hypersurfaces in a Lorentzian manifold.

Arguments in the spirit of those of the proof of Theorem 1.10.1 lead to the construction of *many-body initial data* [46, 47]: starting from initial data for N gravitating isolated systems, one can construct a new initial data set which comprises isometrically compact subsets of each of the original systems, as large as desired, in a distant configuration.

An application of the gluing techniques concerns the question of the existence of CMC slices in space-times with compact Cauchy surfaces. In [15], Bartnik showed that there exist maximally extended, globally hyperbolic solutions of the Einstein equations *with dust* which admit no CMC slices. Later, Eardley and Witt (unpublished) proposed a scheme for showing that similar vacuum solutions exist, but their argument was incomplete. It turns out that these ideas can be implemented using Theorem 1.10.1, which leads to:

Corollary 1.10.2 *[52, 53] There exist maximal globally hyperbolic vacuum space-times with compact Cauchy surfaces which contain no compact spacelike hypersurfaces with constant mean curvature.*

Compact Cauchy surfaces with constant mean curvature are useful objects, as the existence of one such surface gives rise to a unique foliation by such surfaces [30], and hence a canonical choice of time function (often referred to as CMC or York time). Foliations by CMC Cauchy surfaces have also been extensively used in numerical analysis to explore the nature of cosmological singularities. Thus the demonstration that there exist space-times with no such surfaces has a negative impact on such studies.

One natural question is the extent to which space-times with no CMC slices are common among solutions to the vacuum Einstein equations with a fixed spatial topology. It is expected that the examples constructed in [52, 53] are not isolated. In general, there is a great deal of flexibility (in the way of free parameters) in the local gluing construction. This can be used to produce one parameter families of distinct sets of vacuum initial data which lead to space-times as in Corollary 1.10.2. What is less obvious is how to prove that all members of these families give rise to *distinct* maximally extended, globally hyperbolic vacuum space-times.

A deeper question is whether a sequence of space-times which admit constant mean curvature Cauchy surfaces may converge, in a strong topology, to one which admits no such Cauchy surface. (See [12, 15, 77] for general criteria leading to the existence of CMC Cauchy surfaces.)

Non-zero cosmological constant

Gluing constructions have also been carried out with a non-zero cosmological constant [51, 54, 55]. One aim is to construct space-times which coincide, in the asymptotic region, with the corresponding black hole models. In such space-times one has complete control of the geometry in the domain of dependence of the asymptotic region, described there by the Kottler metrics (**??**). For time-symmetric slices of these space-times, the constraint equations reduce to the equation for constant scalar curvature $R = 2\Lambda$. Gluing constructions have been previously carried out in this context, especially in the case of $\Lambda > 0$, but in [51, 54, 55] the emphasis is on gluing with compact support, in the spirit of Corvino's thesis and its extensions already discussed.

The time-symmetric slices of the $\Lambda > 0$ Kottler space-times provide "Delaunay" metrics (see [55] and references therein), and the main result of [54, 55] is the construction of large families of metrics with exactly Delaunay ends. When $\Lambda < 0$ the focus is on asymptotically hyperbolic metrics with constant negative scalar curvature. With hindsight, within the family of Kottler metrics with $\Lambda \in \mathbb{R}$ (with $\Lambda = 0$ corresponding to the Schwarzschild metric), the gluing in the $\Lambda > 0$ setting is technically easiest, while that with $\Lambda < 0$ is the most difficult. This is due to the fact that for $\Lambda > 0$ one deals with one linearized operator with a one-dimensional kernel; in the case $\Lambda = 0$ the kernel is $(n+1)$–dimensional; while for $\Lambda < 0$ one needs to consider a one-parameter family of operators with $(n+1)$–dimensional kernels.

1.11 Other hyperbolic reductions

The wave-coordinates approach of Choquet-Bruhat, presented above, is the first hyperbolic reduction discovered for the Einstein equations. It has been given new life by the Lindblad-Rodnianski stability theorem. However, one should keep in mind the existence of several other such reductions.

An example is given by the symmetric-hyperbolic first order system of Baumgarte, Shapiro, Shibata and Nakamura [21, 143, 150], known as the BSSN system, widely used in numerical general relativity. Another noteworthy example is the elliptic-hyperbolic system of [7], in which the elliptic character of some of the equations provides increased control of the solution. A notorious problem in numerical simulations is the lack of constraint preservation, see [84, 132] and references therein for attempts to improve the situation. The reader is referred to [72, 76] for a review of many other possibilities.

1.12 The characteristic Cauchy problem

Another important systematic construction of solutions of the vacuum Einstein equations proceeds via a *characteristic Cauchy problem.* In this case the initial data are prescribed on Cauchy hypersurfaces which are allowed to be piecewise null. This problem has been considerably less studied than the spacelike one described above. We will not go into any details here; see [27, 31, 32, 37, 44, 62, 141] for further information.

1.13 Initial-boundary value problems

Numerical simulations necessarily take place on a finite grid, which leads to the need of considering initial-boundary value problems. In general relativity those are considerably more complicated than the Cauchy problem, and much remains to be understood. In pioneering work, Friedrich and Nagy [74] constructed a system of equations, equivalent to Einstein's, for a set of fields that includes some components of the Weyl tensor, and proved well-posedness of an initial-boundary value problem for those equations. It would seem that the recent work by Kreiss *et al.* [104] might lead to a simpler formulation of the problem at hand.

Part II

Appendices

Appendix A

Introduction to pseudo-Riemannian geometry

A.1 Manifolds

DEFINITION A.1.1 *An n–dimensional manifold is a set M equipped with the following:*

1. *topology: a "connected Hausdorff paracompact topological space" (think of nicely looking subsets of $\mathbb{R}^{1+n}$, like spheres, hyperboloids, and such), together with*

2. *local charts: a collection of coordinate patches $(\mathscr{U}, x^i)$ covering M, where $\mathscr{U}$ is an open subset of M, with the functions $x^i : \mathscr{U} \to \mathbb{R}^n$ being continuous. One further requires that the maps*

$$M \supset \mathscr{U} \ni p \mapsto (x^1(p), \dots, x^n(p)) \in \mathscr{V} \subset \mathbb{R}^n$$

   are homeomorphisms.

3. *compatibility: given two overlapping coordinate patches, $(\mathscr{U}, x^i)$ and $(\widetilde{\mathscr{U}}, \tilde{x}^i)$, with corresponding sets $\mathscr{V}, \widetilde{\mathscr{V}} \subset \mathbb{R}^n$, the maps $\tilde{x}^j \mapsto x^i(\tilde{x}^j)$ are smooth diffeomorphisms wherever defined: this means that they are bijections differentiable as many times as one wishes, with*

$$\det \left[\frac{\partial x^i}{\partial \tilde{x}^j} \right] \text{ nowhere vanishing} .$$

Definition of differentiability: A function on M is smooth if it is smooth when expressed in terms of local coordinates. Similarly for tensors.

EXAMPLES:

1. $\mathbb{R}^n$ with the usual topology, one single global coordinate patch.

2. A sphere: use stereographic projection to obtain two overlapping coordinate systems (or use spherical angles, but then one must avoid borderline angles, so they don't cover the whole manifold!).

3. We will use several coordinate patches (in fact, five), to describe the Schwarzschild black hole, though one spherical coordinate system would suffice.

4. Let $f : \mathbb{R}^n \to \mathbb{R}$, and define $N := f^{-1}(0)$. If ∇f has no zeros on N, then every connected component of N is a smooth $(n-1)$–dimensional manifold. This construction leads to a plethora of examples. For example, if $f = \sqrt{(x^1)^2 + \ldots + (x^n)^2} - R$, with $R > 0$, then N is a sphere of radius R.

In this context a useful example is provided by the function $f = t^2 - x^2$ on $\mathbb{R}^2$: its zero-level-set is the light-cone $t = \pm x$, which is a manifold except at the origin; note that $\nabla f = 0$ there, which shows that the criterion is sharp.

A.2 Scalar functions

Let M be an n-dimensional manifold. Since manifolds are defined using coordinate charts, we need to understand how things behave under coordinate changes. For instance, under a change of coordinates $x^i \to y^j(x^i)$, to a function $f(x)$ we can associate a new function $\bar{f}(y)$, using the rule

$$\bar{f}(y) = f(x(y)) \quad \Longleftrightarrow \quad f(x) = \bar{f}(y(x)) \ .$$

In general relativity it is a common abuse of notation to write the same symbol f for what we wrote $\bar{f}$, when we think that this is the same function but expressed in a different coordinate system. We then say that a real- or complex-valued f is a *scalar function* when, under a change of coordinates $x \to y(x)$, the function f transforms as $f \to f(x(y))$.

In this section, to make things clearer, we will write $\bar{f}$ for $f(x(y))$ even when f is a scalar, but this will almost never be done in the remainder of these notes. For example we will systematically use the same symbol $g_{\mu\nu}$ for the metric components, whatever the coordinate system used.

A.3 Vector fields

Physicists often think of vector fields in terms of coordinate systems: a vector field X is an object which in a coordinate system $\{x^i\}$ is represented by a collection of functions X^i. In a new coordinate system $\{y^j\}$ the field X is represented by a new set of functions:

$$X^i(x) \to X^j(y) := X^j(x(y)) \frac{\partial y^i}{\partial x^j}(x(y)) \ . \qquad \text{(A.3.1)}$$

(The summation convention is used throughout, so that the index j has to be summed over.)

The notion of a vector field finds its roots in the notion of the tangent to a curve, say $s \to \gamma(s)$. If we use local coordinates to write $\gamma(s)$ as $(\gamma^1(s), \gamma^2(s), \ldots, \gamma^n(s))$, the tangent to that curve at the point $\gamma(s)$ is defined as the set of numbers

$$(\dot{\gamma}^1(s), \dot{\gamma}^2(s), \ldots, \dot{\gamma}^n(s)) \ .$$

Consider, then, a curve $\gamma(s)$ given in a coordinate system x^i and let us perform a change of coordinates $x^i \to y^j(x^i)$. In the new coordinates y^j the curve γ is represented by the functions $y^j(\gamma^i(s))$, with new tangent

$$\frac{dy^j}{ds}(y(\gamma(s))) = \frac{\partial y^j}{\partial x^i}(\gamma(s))\dot\gamma^i(s) .$$

This motivates (A.3.1).

In modern differential geometry a different approach is taken: one identifies vector fields with homogeneous first order differential operators acting on real valued functions $f : M \to \mathbb{R}$. In local coordinates $\{x^i\}$ a vector field X will be written as $X^i\partial_i$, where the X^i's are the "physicists's functions" just mentioned. This means that the action of X on functions is given by the formula

$$\boxed{X(f) := X^i\partial_i f} \qquad \text{(A.3.2)}$$

(recall that ∂_i is the partial derivative with respect to the coordinate x^i). Conversely, given some abstract first order homogeneous derivative operator X, the (perhaps locally defined) functions X^i in (A.3.2) can be found by acting on the coordinate functions:

$$X(x^i) = X^i . \qquad \text{(A.3.3)}$$

One justification for the differential operator approach is the fact that the tangent $\dot\gamma$ to a curve γ can be calculated — in a way independent of the coordinate system $\{x^i\}$ chosen to represent γ — using the equation

$$\dot\gamma(f) := \frac{d(f\circ\gamma)}{dt} .$$

Indeed, if γ is represented as $\gamma(t) = \{x^i = \gamma^i(t)\}$ within a coordinate patch, then we have

$$\frac{d(f\circ\gamma)(t)}{dt} = \frac{d(f(\gamma(t)))}{dt} = \frac{d\gamma^i(t)}{dt}(\partial_i f)(\gamma(t)) ,$$

recovering the previous coordinate formula $\dot\gamma = (d\gamma^i/dt)$.

An even better justification is that *the transformation rule (A.3.1) becomes implicit in the formalism.* Indeed, consider a (scalar) function f, so that the differential operator X acts on f by differentiation:

$$X(f)(x) := \sum_i X^i \frac{\partial f(x)}{\partial x^i} . \qquad \text{(A.3.4)}$$

If we make a coordinate change so that

$$x^j = x^j(y^k) \quad \Longleftrightarrow \quad y^k = y^k(x^j) ,$$

keeping in mind that

$$\bar f(y) = f(x(y)) \quad \Longleftrightarrow \quad f(x) = \bar f(y(x)) ,$$

then

$$\begin{aligned}
X(f)(x) &:= \sum_i X^i(x)\frac{\partial f(x)}{\partial x^i} \\
&= \sum_i X^i(x)\frac{\partial \bar f(y(x))}{\partial x^i} \\
&= \sum_{i,k} X^i(x)\frac{\partial \bar f(y(x))}{\partial y^k}\frac{\partial y^k}{\partial x^i}(x) \\
&= \sum_k \bar X^k(y(x))\frac{\partial \bar f(y(x))}{\partial y^k} \\
&= \left(\sum_k \bar X^k \frac{\partial \bar f}{\partial y^k}\right)(y(x)) \; ,
\end{aligned}$$

with $\bar X^k$ given by the right-hand-side of (A.3.1). So

$X(f)$ is a scalar iff the coefficients X^i satisfy the transformation law of a vector.

EXERCICE A.3.1 Check that this is a necessary and sufficient condition.

One often uses the middle formula in the above calculation in the form

$$\frac{\partial}{\partial x^i} = \frac{\partial y^k}{\partial x^i}\frac{\partial}{\partial y^k} \; . \tag{A.3.5}$$

Note that the tangent to the curve $s \to (s, x^2, x^3, \dots x^n)$, where $(x^2, x^3, \dots x^n)$ are constants, is identified with the differential operator

$$\partial_1 \equiv \frac{\partial}{\partial x^1} \; .$$

Similarly the tangent to the curve $s \to (x^1, s, x^3, \dots x^n)$, where $(x^1, x^3, \dots x^n)$ are constants, is identified with

$$\partial_2 \equiv \frac{\partial}{\partial x^2} \; ,$$

etc. Thus, $\dot\gamma$ is identified with

$$\dot\gamma(s) = \dot\gamma^i \partial_i$$

At any given point $p \in M$ the set of vectors forms a vector space, denoted by T_pM. The collection of all the tangent spaces is called the tangent bundle to M, denoted by TM.

A.3.1 Lie bracket

Vector fields can be added and multiplied by functions in the obvious way. Another useful operation is the *Lie bracket*, or *commutator*, defined as

$$\boxed{[X,Y](f) := X(Y(f)) - Y(X(f))}\ . \tag{A.3.6}$$

One needs to check that this does indeed define a new vector field: the simplest way is to use local coordinates,

$$\begin{aligned}
[X,Y](f) &= X^j\partial_j(Y^i\partial_i f) - Y^j\partial_j(X^i\partial_i f)\\
&= X^j(\partial_j(Y^i)\partial_i f + Y^i\partial_j\partial_i f) - Y^j(\partial_j(X^i)\partial_i f + X^i\partial_j\partial_i f)\\
&= (X^j\partial_j Y^i - Y^j\partial_j X^i)\partial_i f + \underbrace{X^jY^i\partial_j\partial_i f - Y^jX^i\partial_j\partial_i f}_{=X^jY^i\underbrace{(\partial_j\partial_i f - \partial_i\partial_j f)}_{0}}\\
&= (X^j\partial_j Y^i - Y^j\partial_j X^i)\partial_i f\ ,
\end{aligned} \tag{A.3.7}$$

which is indeed a homogeneous first order differential operator. Here we have used the symmetry of the matrix of second derivatives of twice differentiable functions. We note that the last line of (A.3.7) also gives an explicit coordinate expression for the commutator of two differentiable vector fields.

The Lie bracket satisfies the *Jacobi identity*:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0\ .$$

Indeed, if we write $S_{X,Y,Z}$ for a cyclic sum, then

$$\begin{aligned}
&([X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]])(f) = S_{X,Y,Z}[X,[Y,Z]](f)\\
&\quad = S_{X,Y,Z}\left\{X([Y,Z](f)) - [Y,Z](X(f))\right\}\\
&\quad = S_{X,Y,Z}\left\{X(Y(Z(f))) - X(Z(Y(f))) - Y(Z(X(f))) + Z(Y(X(f)))\right\}\ .
\end{aligned}$$

The third term is a cyclic permutation of the first, and the fourth a cyclic permutation of the second, so the sum gives zero.

A.4 Covectors

Covectors are *maps from the space of vectors to functions which are linear under addition and multiplication by functions.*

The basic object is the *coordinate differential* dx^i, defined by its action on vectors as follows:

$$dx^i(X^j\partial_j) := X^i\ . \tag{A.4.1}$$

Equivalently,

$$dx^i(\partial_j) := \delta^i_j := \begin{cases} 1, & i = j;\\ 0, & \text{otherwise.}\end{cases}$$

The dx^i's form a basis for the space of covectors: indeed, let φ be a linear map on the space of vectors, then

$$\varphi(\underbrace{X}_{X^i\partial_i}) = \varphi(X^i\partial_i) \underbrace{=}_{\text{linearity}} X^i \underbrace{\varphi(\partial_i)}_{\text{call this } \varphi_i} = \varphi_i dx^i(X) \underbrace{=}_{\text{def. of sum of functions}} (\varphi_i dx^i)(X)\ ,$$

hence

$$\varphi = \varphi_i dx^i ,$$

and every φ can indeed be written as a linear combination of the dx^i's. Under a change of coordinates we have

$$\bar\varphi_i \bar X^i = \bar\varphi_i \frac{\partial y^i}{\partial x^k} X^k = \varphi_k X^k ,$$

leading to the following transformation law for components of covectors:

$$\varphi_k = \bar\varphi_i \frac{\partial y^i}{\partial x^k} , \tag{A.4.2}$$

Given a scalar f, we define its *differential* df as

$$df = \frac{\partial f}{\partial x^1} dx^1 + \ldots + \frac{\partial f}{\partial x^n} dx^n .$$

With this definition, dx^i *is* the differential of the coordinate function x^i.

As presented above, the differential of a function is a covector by definition. As an exercice, you should check directly that the collection of functions $\varphi_i := \partial_i f$ satisfies the transformation rule (A.4.2).

We have a formula which is often used in calculations

$$dy^j = \frac{\partial y^j}{\partial x^k} dx^k .$$

An elegant approach to the definition of differentials proceeds as follows: Given any function f, we define:

$$df(X) := X(f) . \tag{A.4.3}$$

(Recall that here we are viewing a vector field X as a differential operator on functions, defined by (A.3.4).) The map $X \mapsto df(X)$ is linear under addition of vectors, and multiplication of vectors by numbers: if λ, μ are real numbers, and X and Y are vector fields, then

$$\begin{aligned} df(\lambda X + \mu Y) & \underbrace{=}_{\text{by definition (A.4.3)}} (\lambda X + \mu Y)(f) \\ & \underbrace{=}_{\text{by definition (A.3.4)}} \lambda X^i \partial_i f + \mu Y^i \partial_i f \\ & \underbrace{=}_{\text{by definition (A.4.3)}} \lambda df(X) + \mu df(Y) . \end{aligned}$$

Applying (A.4.3) to the function $f = x^i$ we obtain

$$dx^i(\partial_j) = \frac{\partial x^i}{\partial x^j} = \delta^i_j ,$$

recovering (A.4.1).

EXAMPLE A.4.1 Let (ρ, φ) be polar coordinates on $\mathbb{R}^2$, thus $x = \rho\cos\varphi$, $y = \rho\sin\varphi$, and then

$$\begin{aligned} dx &= d(\rho\cos\varphi) = \cos\varphi d\rho - \rho\sin\varphi d\varphi \ , \\ dy &= d(\rho\sin\varphi) = \sin\varphi d\rho + \rho\cos\varphi d\varphi \ . \end{aligned}$$

At any given point $p \in M$, the set of covectors forms a vector space, denoted by T_p^*M. The collection of all the tangent spaces is called the cotangent bundle to M, denoted by T^*M.

Summarising, covectors are dual to vectors. It is convenient to define

$$\boxed{dx^i(X) := X^i} \ ,$$

where X^i is as in (A.3.2). With this definition the (locally defined) bases $\{\partial_i\}_{i=1,\ldots,\dim M}$ of TM and $\{dx^j\}_{i=1,\ldots,\dim M}$ of T^*M are dual to each other:

$$\langle dx^i, \partial_j\rangle := dx^i(\partial_j) = \delta^i_j \ ,$$

where δ^i_j is the Kronecker delta, equal to one when $i = j$ and zero otherwise.

A.5 Bilinear maps, two-covariant tensors

A map is said to be multi-linear if it is linear in every entry; e.g. g is bilinear if

$$g(aX + bY, Z) = ag(X, Z) + bg(Y, Z) \ ,$$

and

$$g(X, aZ + bW) = ag(X, Z) + bg(X, W) \ .$$

Here, as elsewhere when talking about *tensors*, bilinearity is meant with respect to addition and to multiplication by functions.

A map g which is bilinear on the space of vectors can be represented by a matrix with two indices down:

$$g(X, Y) = g(X^i\partial_i, Y^j\partial_j) = X^iY^j \underbrace{g(\partial_i, \partial_j)}_{=:g_{ij}} = g_{ij}X^iY^j = g_{ij}dx^i(X)dx^j(Y) \ .$$

We say that g is a *covariant tensor of valence two.*

We say that g is *symmetric* if $g(X, Y) = g(Y, X)$ for all X, Y; equivalently, $g_{ij} = g_{ji}$.

A symmetric bilinear tensor field is said to be *non-degenerate* if $\det g_{ij}$ has no zeros.

By Sylvester's inertia theorem, there exists a basis θ^i of the space of covectors so that a symmetric bilinear map g can be written as

$$g(X,Y) = \theta^1(X)\theta^1(Y) + \ldots + \theta^s(X)\theta^s(Y) - \theta^{s+1}(X)\theta^{s+1}(Y) - \ldots - \theta^{s+r}(X)\theta^{s+r}(Y)$$

(s, r) is called the signature of g; in geometry, unless specifically said otherwise, one always assumed that the signature does not change from point to point.

If $s = n$, in dimension n, then g is said to be a Riemannian metric tensor.

A canonical example is provided by the flat Riemannian metric on $\mathbb{R}^n$ is

$$g = (dx^1)^2 + \ldots + (dx^n)^2 \ .$$

By definition, a *Riemannian metric* is a field of symmetric two-covariant tensors with signature $(+, \ldots, +)$ and with $\det g_{ij}$ without zeros.

A Riemannian metric can be used to define the length of curves: if $\gamma : [a, b] \ni s \to \gamma(s)$, then

$$\ell_g(\gamma) = \int_a^b \sqrt{g(\dot\gamma, \dot\gamma)} ds \ .$$

One can then define the distance between points by minimizing the length of the curves connecting them.

If $s = 1$ and $r = N - 1$, in dimension N, then g is said to be a *Lorentzian metric tensor*.

For example, the *Minkowski metric* on $\mathbb{R}^{1+n}$ is

$$\eta = (dx^0)^2 - (dx^1)^2 - \ldots - (dx^n)^2 \ .$$

A.6 Tensor products

If φ and θ are covectors we can define a bilinear map using the formula

$$(\varphi \otimes \theta)(X, Y) = \varphi(X)\theta(Y) \ . \tag{A.6.1}$$

For example

$$(dx^1 \otimes dx^2)(X, Y) = X^1 Y^2 \ .$$

Using this notation we have

$$g(X, Y) = g(X^i \partial_i, Y^j \partial_j) = \underbrace{g(\partial_j, \partial_j)}_{=:g_{ij}} \underbrace{\underbrace{X^i}_{dx^i(X)} \underbrace{Y^j}_{dx^j(Y)}}_{(dx^i \otimes dx^j(X,Y)} = (g_{ij} dx^i \otimes dx^j)(X, Y)$$

We will write $dx^i dx^j$ for the symmetric product,

$$dx^i dx^j := \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i) \ ,$$

and $dx^i \wedge dx^j$ for the anti-symmetric one,

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

It should be clear how this generalises: the tensors $dx^i \otimes dx^j \otimes dx^k$, defined as

$$(dx^i \otimes dx^j \otimes dx^k)(X, Y, Z) = X^i Y^j Z^k \ ,$$

form a basis of three-linear maps on the space of vectors, which are objects of the form

$$X = X_{ijk} dx^i \otimes dx^j \otimes dx^k \ .$$

Here X is a called *tensor of valence* $(0,3)$. Each index transforms as for a covector:

$$X = X_{ijk} dx^i \otimes dx^j \otimes dx^k = X_{ijk} \frac{\partial x^i}{\partial y^m} \frac{\partial x^j}{\partial y^\ell} \frac{\partial x^k}{\partial y^n} dy^m \otimes dy^\ell \otimes dy^n \ .$$

It is sometimes useful to think of vectors as linear maps on co-vectors, using a formula which looks funny when first met: if θ is a covector, and X is a vector, then

$$X(\theta) := \theta(X) \ .$$

So if $\theta = \theta_i dx^i$ and $X = X^i \partial_i$ then

$$\theta(X) = \theta_i X^i = X^i \theta_i = X(\theta) \ .$$

It then makes sense to define e.g. $\partial_i \otimes \partial_j$ as a bilinear map on covectors:

$$(\partial_i \otimes \partial_j)(\theta, \psi) := \theta_i \psi_j \ .$$

And one can define a map $\partial_i \otimes dx^j$ which is linear on forms in the first slot, and linear in vectors in the second slot as

$$(\partial_i \otimes dx^j)(\theta, X) := \partial_i(\theta) dx^j(X) = \theta_i X^j \ . \tag{A.6.2}$$

The $\partial_i \otimes dx^j$'s form the basis of the space of *tensors of rank* $(1,1)$:

$$T = T^i{}_j \partial_i \otimes dx^j \ .$$

Generally, a *tensor of valence, or rank,* (r,s) can be defined as an object which has r vector indices and s covector indices, so that it transforms as

$$S^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \to S^{m_1 \ldots m_r}{}_{\ell_1 \ldots \ell_s} \frac{\partial y^{i_1}}{\partial x^{m_1}} \cdots \frac{\partial y^{i_s}}{\partial x^{m_r}} \frac{\partial x^{\ell_1}}{\partial y^{j_1}} \cdots \frac{\partial x^{\ell_s}}{\partial y^{j_s}}$$

For example, if $X = X^i \partial_i$ and $Y = Y^j \partial_j$ are vectors, then $X \otimes Y = X^i Y^j \partial_i \otimes \partial_j$ forms a contravariant tensor of valence two.

Tensors of same valence can be added in the obvious way: *e.g.*

$$(A+B)(X,Y) := A(X,Y) + B(X,Y) \quad \Longleftrightarrow \quad (A+B)_{ij} = A_{ij} + B_{ij} \ .$$

Tensors can be multiplied by scalars: *e.g.*

$$(fA)(X,Y,Z) := fA(X,Y,Z) \quad \Longleftrightarrow \quad f(A_{ijk}) := (fA_{ijk}) \ .$$

Finally, we have seen in (A.6.1) how to take tensor products for one forms, and in (A.6.2) how to take a tensor product of a vector and a one form, but this can also be done for higher order tensor; e.g., if S is of valence (a,b) and T is a multilinear map of valence (c,d), then $S \otimes T$ is a multilinear map of valence $(a+c, b+d)$, defined as

$$(S \otimes T)(\underbrace{\theta, \ldots}_{a \text{ covectors and } b \text{ vectors}}, \underbrace{\psi, \ldots}_{c \text{ covectors and } d \text{ vectors}}) := S(\theta, \ldots) T(\psi, \ldots) \ .$$

A.6.1 Contractions

Given a tensor field $S^i{}_j$ with one index down and one index up one can perform the sum

$$S^i{}_i \ .$$

This defines a scalar, i.e., a function on the manifold. Indeed, using the transformation rule

$$S^i{}_j \to \bar{S}^\ell{}_k = S^i{}_j \frac{\partial x^j}{\partial y^k} \frac{\partial y^\ell}{\partial x^i} \ ,$$

one finds

$$\bar{S}^\ell{}_\ell = S^i{}_j \underbrace{\frac{\partial x^j}{\partial y^\ell} \frac{\partial y^\ell}{\partial x^i}}_{\delta^j_i} = S^i{}_i \ ,$$

as desired.

One can similarly do contractions on higher valence tensors, e.g.

$$S^{i_1 i_2 \ldots i_r}{}_{j_1 j_2 j_3 \ldots j_s} \to S^{\ell i_2 \ldots i_r}{}_{j_1 \ell j_3 \ldots j_s} \ .$$

After contraction, a tensor of rank $(r+1, s+1)$ becomes of rank (r, s).

A.7 Raising and lowering of indices

Let g be a symmetric two-covariant tensor field on M, by definition such an object is the assignment to each point $p \in M$ of a bilinear map $g(p)$ from $T_pM \times T_pM$ to $\mathbb{R}$, with the additional property

$$g(X, Y) = g(Y, X) \ .$$

In this work the symbol g will be reserved to *non-degenerate* symmetric two-covariant tensor fields. It is usual to simply write g for $g(p)$, the point p being implicitly understood. We will sometimes write g_p for $g(p)$ when referencing p will be useful.

The usual Sylvester's inertia theorem tells us that at each p the map g will have a well defined signature; clearly this signature will be point-independent on a connected manifold when g is non-degenerate. A pair (M, g) is said to be a *Riemannian manifold* when the signature of g is $(\dim M, 0)$; equivalently, when g is a positive definite bilinear form on every product $T_pM \times T_pM$. A pair (M, g) is said to be a *Lorentzian manifold* when the signature of g is $(\dim M - 1, 1)$. One talks about *pseudo-Riemannian* manifolds whatever the signature of g, as long as g is non-degenerate, but we will only encounter Riemannian and Lorentzian metrics in this work.

Since g is non-degenerate it induces an isomorphism

$$\flat : T_pM \to T_p^*M$$

by the formula

$$\boxed{X_\flat(Y) = g(X, Y)} \ .$$

In local coordinates this gives

$$X_\flat = g_{ij} X^i dx^j =: X_j dx^j \ . \tag{A.7.1}$$

This last equality defines X_j — "the vector X^j with the index j lowered":

$$\boxed{X_i := g_{ij} X^j} \ . \tag{A.7.2}$$

The operation (A.7.2) is called the *lowering of indices* in the physics literature and, again in the physics literature, one does not make a distinction between the one-form $X_\flat$ and the vector X.

The inverse map will be denoted by $\sharp$ and is called the *raising of indices*; from (A.7.1) we obviously have

$$\alpha^\sharp = g^{ij} \alpha_i \partial_j =: \alpha^i \partial_i \quad \Longleftrightarrow \quad dx^i(\alpha^\sharp) = \boxed{\alpha^i = g^{ij} \alpha_j} \ ,$$

where g^{ij} is the matrix inverse to g_{ij}. For example,

$$(dx^i)^\sharp = g^{ik} \partial_k \ .$$

Clearly g^{ij}, understood as the matrix of a bilinear form on $T^*_p M$, has the same signature as g, and can be used to define a scalar product $g^\sharp$ on $T^*_p(M)$:

$$g^\sharp(\alpha, \beta) := g(\alpha^\sharp, \beta^\sharp) \quad \Longleftrightarrow \quad g^\sharp(dx^i, dx^j) = g^{ij} \ .$$

This last equality is justified as follows:

$$g^\sharp(dx^i, dx^j) = g((dx^i)^\sharp, (dx^j)^\sharp) = g(g^{ik} \partial_k, g^{j\ell} \partial_\ell) = \underbrace{g^{ik} g_{k\ell}}_{=\delta^i_\ell} g^{j\ell} = g^{ji} = g^{ij} \ .$$

It is convenient to use the same letter g for $g^\sharp$ — physicists do it all the time — or for scalar products induced by g on all the remaining tensor bundles, and we will sometimes do so.

A.8 The Lie derivative

We start with a pedestrian approach to the definition of Lie derivative; the elegant geometric definition is given at the end of the section.

Given a vector field X, the *Lie derivative* $\mathscr{L}_X$ is an operation on tensor fields, defined as follows:

For a function f, one sets

$$\mathscr{L}_X f := X(f) \ . \tag{A.8.1}$$

For a vector field Y, the Lie derivative coincides with the Lie bracket:

$$\mathscr{L}_X Y := [X, Y] \ . \tag{A.8.2}$$

For a one form α, $\mathscr{L}_X\alpha$ is defined by imposing the Leibniz rule written backwards:

$$(\mathscr{L}_X\alpha)(Y) := \mathscr{L}_X(\alpha(Y)) - \alpha(\mathscr{L}_X Y) \;. \tag{A.8.3}$$

(Indeed, the Leibniz rule applied to the contraction $\alpha_i X^i$ would read

$$\mathscr{L}_X(\alpha_i Y^i) = (\mathscr{L}_X\alpha)_i Y^i + \alpha_i(\mathscr{L}_X Y)^i \;,$$

which can be rewritten as (A.8.3).)

Let us check that (A.8.3) defines a one form. Clearly, the right-hand side transforms in the desired way when Y is replaced by Y_1+Y_2. Now, if we replace Y by fY, where f is a function, then

$$\begin{aligned}
(\mathscr{L}_X\alpha)(fY) &= \mathscr{L}_X(\alpha(fY)) - \alpha(\underbrace{\mathscr{L}_X(fY)}_{X(f)Y+f\mathscr{L}_X Y})\\
&= X(f\alpha(Y)) - \alpha(X(f)Y + f\mathscr{L}_X Y))\\
&= X(f)\alpha(Y) + fX(\alpha(Y)) - \alpha(X(f)Y) - \alpha(f\mathscr{L}_X Y))\\
&= fX(\alpha(Y)) - f\alpha(\mathscr{L}_X Y))\\
&= f((\mathscr{L}_X\alpha)(Y)) \;.
\end{aligned}$$

So $\mathscr{L}_X\alpha$ is a C^∞-linear map on vector fields, hence a covector field.

In coordinate-components notation we have

$$(\mathscr{L}_X\alpha)_a = X^b\partial_b\alpha_a + \alpha_b\partial_a X^b \;.$$

Indeed,

$$\begin{aligned}
(\mathscr{L}_X\alpha)_i Y^i &:= \mathscr{L}_X(\alpha_i Y^i) - \alpha_i(\mathscr{L}_X Y)^i\\
&= X^k\partial_k(\alpha_i Y^i) - \alpha_i(X^k\partial_k Y^i - Y^k\partial_k X^i)\\
&= X^k(\partial_k\alpha_i)Y^i + \alpha_i Y^k\partial_k X^i\\
&= \left(X^k\partial_k\alpha_i + \alpha_k\partial_i X^k\right)Y^i \;,
\end{aligned}$$

as desired

For tensor products, the Lie derivative is defined by imposing linearity under addition together with the Leibniz rule:

$$\mathscr{L}_X(\alpha\otimes\beta) = (\mathscr{L}_X\alpha)\otimes\beta + \alpha\otimes\mathscr{L}_X\beta \;.$$

Since a general tensor A is a sum of tensor products,

$$A = A^{a_1\ldots a_p}{}_{b_1\ldots b_q}\partial_{a_1}\otimes\ldots\partial_{a_p}\otimes dx^{b_1}\otimes\ldots\otimes dx^{a_p} \;,$$

requiring linearity with respect to addition of tensors gives thus a definition of Lie derivative for any tensor.

For example, we claim that

$$\mathscr{L}_X T^a{}_b = X^c\partial_c T^a{}_b - T^c{}_b\partial_c X^a + T^a{}_c\partial_b X^c \;, \tag{A.8.4}$$

To see this, call a tensor $T^a{}_b$ *simple* if it is of the form $Y \otimes \alpha$, where Y is a vector and α is a covector. Using indices, this corresponds to $Y^a\alpha_b$ and so, by the Leibniz rule,

$$\begin{aligned}
\mathscr{L}_X(Y\otimes\alpha)^a{}_b &= \mathscr{L}_X(Y^a\alpha_b)\\
&= (\mathscr{L}_XY)^a\alpha_b + Y^a(\mathscr{L}_X\alpha)_b\\
&= (X^c\partial_cY^a - Y^c\partial_cX^a)\alpha_b + Y^a(X^c\partial_c\alpha_b + \alpha_c\partial_bX^c)\\
&= X^c\partial_c(Y^a\alpha_b) - Y^c\alpha_b\partial_cX^a + Y^a\alpha_c\partial_bX^c\ ,
\end{aligned}$$

which coincides with (A.8.4) if $T^a{}_b = Y^b\alpha_b$. But a general $T^a{}_b$ can be written as a linear combination with constant coefficients of simple tensors,

$$T = \sum_{a,b} \underbrace{T^a{}_b\partial_a \otimes dx^b}_{\text{no summation, so simple}}\ ,$$

and the result follows.

Similarly, one has, e.g.,

$$\mathscr{L}_XR^{ab} = X^c\partial_cR^{ab} - R^{ac}\partial_cX^b - R^{bc}\partial_cX^a\ ,$$

$$\mathscr{L}_XS_{ab} = X^c\partial_cS_{ab} + S_{ac}\partial_bX^c + S_{bc}\partial_aX^c\ , \qquad \text{(A.8.5)}$$

etc. Those are all special cases of the general formula for the Lie derivative $\mathscr{L}_XA^{a_1\ldots a_p}{}_{b_1\ldots b_q}$:

$$\begin{aligned}
\mathscr{L}_XA^{a_1\ldots a_p}{}_{b_1\ldots b_q} &= X^c\partial_cA^{a_1\ldots a_p}{}_{b_1\ldots b_q} - A^{ca_2\ldots a_p}{}_{b_1\ldots b_q}\partial_cX^{a_1} - \ldots\\
&\quad + A^{a_1\ldots a_p}{}_{cb_1\ldots b_q}\partial_{b_1}X^c + \ldots\ .
\end{aligned}$$

A useful property of Lie derivatives is

$$\mathscr{L}_{[X,Y]} = [\mathscr{L}_X, \mathscr{L}_Y]\ , \qquad \text{(A.8.6)}$$

where, for a tensor T, the commutator $[\mathscr{L}_X, \mathscr{L}_Y]T$ is defined in the usual way:

$$[\mathscr{L}_X, \mathscr{L}_Y]T := \mathscr{L}_X(\mathscr{L}_YT) - \mathscr{L}_Y(\mathscr{L}_XT)\ . \qquad \text{(A.8.7)}$$

To see this, we first note that if $T = f$ is a function, then the right-hand-side of (A.8.7) is the definition of $[X,Y](f)$, which in turn coincides with the definition of $\mathscr{L}_{[X,Y]}(f)$.

Next, for a vector field $T = Z$, (A.8.6) reads

$$\mathscr{L}_{[X,Y]}Z = \mathscr{L}_X(\mathscr{L}_YZ) - \mathscr{L}_Y(\mathscr{L}_XZ)\ , \qquad \text{(A.8.8)}$$

which is the same as

$$[[X,Y],Z] = [X,[Y,Z]] - [Y,[X,Z]]\ , \qquad \text{(A.8.9)}$$

which is the same as

$$[Z,[Y,X]] + [X,[Z,Y]] + [Y,[X,Z]] = 0\ , \qquad \text{(A.8.10)}$$

which is the Jacobi identity. Hence (A.8.6) holds for vector fields.

We continue with a one form α. We use the definitions, with Z any vector field:

$$\begin{aligned}(\mathscr{L}_X\mathscr{L}_Y\alpha)(Z) &= X(\underbrace{(\mathscr{L}_Y\alpha)(Z)}_{Y(\alpha(Z))-\alpha(\mathscr{L}_Y Z))}) - \underbrace{(\mathscr{L}_Y\alpha)(\mathscr{L}_X Z)}_{Y(\alpha(\mathscr{L}_X Z))-\alpha(\mathscr{L}_Y\mathscr{L}_X Z)} \\ &= X(Y(\alpha(Z))) - X(\alpha(\mathscr{L}_Y Z))) - Y(\alpha(\mathscr{L}_X Z)) + \alpha(\mathscr{L}_Y\mathscr{L}_X Z)\ .\end{aligned}$$

Antisymmetrizing over X and Y, the second and third term above cancel out, so that

$$\begin{aligned}((\mathscr{L}_X\mathscr{L}_Y\alpha - \mathscr{L}_Y\mathscr{L}_X)\alpha)(Z) &= X(Y(\alpha(Z))) + \alpha(\mathscr{L}_Y\mathscr{L}_X Z) - (X \longleftrightarrow Y) \\ &= [X,Y](\alpha(Z)) - \alpha(\mathscr{L}_X\mathscr{L}_Y Z - \mathscr{L}_Y\mathscr{L}_X Z) \\ &= \mathscr{L}_{[X,Y]}(\alpha(Z)) - \alpha(\mathscr{L}_{[X,Y]}Z) \\ &= (\mathscr{L}_{[X,Y]}\alpha)(Z)\ .\end{aligned}$$

Since Z is arbitrary, (A.8.6) for covectors follows.

To conclude that (A.8.6) holds for arbitrary tensor fields, we note that by construction we have

$$\mathscr{L}_{[X,Y]}(A\otimes B) = \mathscr{L}_{[X,Y]}A\otimes B + A\otimes\mathscr{L}_{[X,Y]}B\ . \tag{A.8.11}$$

Similarly

$$\begin{aligned}\mathscr{L}_X\mathscr{L}_Y(A\otimes B) &= \mathscr{L}_X(\mathscr{L}_Y A\otimes B + A\otimes\mathscr{L}_Y B) \\ &= \mathscr{L}_X\mathscr{L}_Y A\otimes B + \mathscr{L}_X A\otimes\mathscr{L}_Y B + \mathscr{L}_Y A\otimes\mathscr{L}_X B \\ &\quad + A\otimes\mathscr{L}_X\mathscr{L}_Y B\ .\end{aligned} \tag{A.8.12}$$

Exchanging X with Y and subtracting, the middle terms drop out:

$$[\mathscr{L}_X,\mathscr{L}_Y](A\otimes B) = [\mathscr{L}_X,\mathscr{L}_Y]A\otimes B + A\otimes[\mathscr{L}_X,\mathscr{L}_Y]B\ . \tag{A.8.13}$$

Basing on what has been said, the reader should have no difficulties finishing the proof of (A.8.6).

EXAMPLE A.8.1 As an example of application of the formalism, suppose that there exists a coordinate system in which $(X^a) = (1,0,0,0)$ and $\partial_0 g_{bc} = 0$. Then

$$\mathscr{L}_X g_{ab} = \partial_0 g_{ab} = 0\ .$$

But the Lie derivative of a tensor field is a tensor field, and we conclude that $\mathscr{L}_X g_{ab} = 0$ holds in every coordinate system.

Vector fields for which $\mathscr{L}_X g_{ab} = 0$ are called *Killing vectors*: they arise from symmetries of space-time. We have the useful formula

$$\mathscr{L}_X g_{ab} = \nabla_a X_b + \nabla_b X_a\ . \tag{A.8.14}$$

An effortless proof of this proceeds as follows: in adapted coordinates in which the derivatives of the metric vanish at a point p, one immediately checks that equality holds at p. But both sides are tensor fields, therefore the result holds at p for all coordinate systems, and hence also everywhere.

The brute-force proof of (A.8.14) proceeds as follows:

$$\begin{aligned}
\mathscr{L}_X g_{ab} &= X^c\partial_c g_{ab} + \partial_a X^c g_{cb} + \partial_b X^c g_{ca} \\
&= X^c\partial_c g_{ab} + \partial_a (X^c g_{cb}) - X^c\partial_a g_{cb} + \partial_b (X^c g_{ca}) - X^c\partial_b g_{ca} \\
&= \partial_a X_b + \partial_b X_a + X^c \underbrace{(\partial_c g_{ab} - \partial_a g_{cb} - \partial_b g_{ca})}_{-2g_{cd}\Gamma^d_{ab}} \\
&= \nabla_a X_b + \nabla_b X_a \; .
\end{aligned}$$

A.9 Covariant derivatives

When dealing with $\mathbb{R}^n$, or subsets thereof, there exists an obvious prescription for how to differentiate tensor fields: in this case we have at our disposal the canonical "trivialization $\{\partial_i\}_{i=1,\ldots,n}$ of $T\mathbb{R}^n$" (this means: a globally defined set of vectors which, at every point, form a basis of the tangent space), together with its dual trivialization $\{dx^j\}_{i=1,\ldots,n}$ of $T^*\mathbb{R}^n$. We can expand a tensor field T of valence (k,ℓ) in terms of those bases,

$$\begin{aligned}
&T = T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell}\partial_{i_1}\otimes\ldots\otimes\partial_{i_k}\otimes dx^{j_1}\otimes\ldots\otimes dx^{j_\ell} \\
&\iff \quad T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell} = T(dx^{i_1},\ldots,dx^{i_k},\partial_{j_1},\ldots,\partial_{j_\ell}) \; , \qquad \text{(A.9.1)}
\end{aligned}$$

and differentiate each component $T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell}$ of T separately:

$$X(T)_{\text{in the coordinate system } x^i} := X^i\frac{\partial T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell}}{\partial x^i}\partial_{x^{i_1}}\otimes\ldots\otimes\partial_{x^{i_k}}\otimes dx^{j_1}\otimes\ldots\otimes dx^{j_\ell} \; . \qquad \text{(A.9.2)}$$

The resulting object does, however, *not* behave as a tensor under coordinate transformations, in the sense that the above form of the right-hand-side will *not* be preserved under coordinate transformations: as an example, consider the one-form $T = dx$ on $\mathbb{R}^n$, which has vanishing derivative as defined by (A.9.2). When expressed in spherical coordinates we have

$$T = d(\rho\cos\varphi) = -\rho\sin\varphi d\varphi + \cos\varphi d\rho \; ,$$

the partial derivatives of which are non-zero (both with respect to the original cartesian coordinates (x,y) and to the new spherical ones (ρ,φ)).

The Lie derivative $\mathscr{L}_X$ of Section A.8 maps tensors to tensors but does not resolve this question, because it is *not* linear under multiplication of X by a function.

The notion of *covariant derivative*, sometimes also referred to as *connection*, is introduced precisely to obtain a notion of derivative which has tensorial properties. By definition, a covariant derivative is a map which to a vector field X and a tensor field T assigns a tensor field of the same type as T, denoted by $\nabla_X T$, with the following properties:

1. $\nabla_X T$ is linear with respect to addition both with respect to X and T:

$$\nabla_{X+Y}T = \nabla_X T + \nabla_Y T \; , \qquad \nabla_X(T+Y) = \nabla_X T + \nabla_X Y \; ; \qquad \text{(A.9.3)}$$

2. $\nabla_X T$ is linear with respect to multiplication of X by functions f,

$$\nabla_{fX} T = f \nabla_X T \; ; \tag{A.9.4}$$

3. and, finally, $\nabla_X T$ satisfies the *Leibniz rule* under multiplication of T by a differentiable function f:

$$\nabla_X (fT) = f \nabla_X T + X(f) T \; . \tag{A.9.5}$$

By definition, if T is a tensor field of rank (p, q), then for any vector field X the field $\nabla_X T$ is again a tensor of type (p, q). Since $\nabla_X T$ is linear in X, the field ∇T can naturally be viewed as a tensor field of rank $(p, q + 1)$.

It is natural to ask whether covariant derivatives do exist at all in general and, if so, how many of them can there be. First, it immediately follows from the axioms above that if D and ∇ are two covariant derivatives, then

$$\Delta(X, T) := D_X T - \nabla_X T$$

is multi-linear both with respect to addition and multiplication by functions — the non-homogeneous terms $X(f)T$ in (A.9.5) cancel — and is thus a tensor field. Reciprocally, if ∇ is a covariant derivative and $\Delta(X, T)$ is bilinear with respect to addition and multiplication by functions, then

$$D_X T := \nabla_X T + \Delta(X, T) \tag{A.9.6}$$

is a new covariant derivative. So, at least locally, on tensors of valence (r, s) there are as many covariant derivatives as tensors of valence $(r + s, r + s + 1)$.

We note that the sum of two covariant derivatives is *not* a covariant derivative. However, *convex* combinations of covariant derivatives, with coefficients which may vary from point to point, are again covariant derivatives. This remark allows one to construct covariant derivatives using partitions of unity: Let, indeed, $\{\mathscr{O}_i\}_{i\in\mathbb{N}}$ be an open covering of M by coordinate patches and let φ_i be an associated partition of unity. In each of those coordinate patches we can decompose a tensor field T as in (A.9.1), and define

$$D_X T := \sum_i \varphi_i X^j \partial_j (T^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell}) \partial_{i_1} \otimes \ldots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_\ell} \; . \tag{A.9.7}$$

This procedure, which depends upon the choice of the coordinate patches and the choice of the partition of unity, defines *one* covariant derivative; all other covariant derivatives are then obtained from D using (A.9.6). Note that (A.9.2) is a special case of (A.9.7) when there exists a global coordinate system on M. Thus (A.9.2) *does* define a covariant derivative. However, the associated operation on tensor fields will *not* take the simple form (A.9.2) when we go to a different coordinate system $\{y^i\}$ in general.

A.9.1 Functions

The *canonical covariant derivative on functions* is defined as

$$\nabla_X(f) = X(f) \;,$$

and we will always use the above. This has all the right properties, so obviously covariant derivatives of functions exist. From what has been said, any covariant derivative on functions is of the form

$$\nabla_X f = X(f) + \alpha(X) f \;, \tag{A.9.8}$$

where α is a one-form. Conversely, given any one form α, (A.9.8) defines a covariant derivative on functions. The addition of the lower-order term $\alpha(X)f$ (A.9.8) does not appear to be very useful here, but it turns out to be useful in geometric formulation of electrodynamics, or in *geometric quantization*. In any case such lower-order terms play an essential role when defining covariant derivatives of tensor fields.

A.9.2 Vectors

The simplest next possibility is that of a covariant derivative of vector fields. Let us not worry about existence at this stage, but assume that a covariant derivative exists, and work from there. (Anticipating, we will show shortly that a metric defines a covariant derivative, called the *Levi-Civita* covariant derivative, which is the unique covariant derivative operator satisfying a natural set of conditions, to be discussed below.)

We will first assume that we are working on a set $\Omega \subset M$ over which we have a *global trivialization* of the tangent bundle TM; by definition, this means that there exist vector fields e_a, $a = 1, \ldots, \dim M$, such that at every point $p \in \Omega$ the fields $e_a(p) \in T_pM$ form a basis of T_pM.[1]

Let θ^a denote the dual trivialization of T^*M — by definition the θ^a's satisfy

$$\boxed{\theta^a(e_b) = \delta^a_b} \;.$$

Given a covariant derivative ∇ on vector fields we set

$$\Gamma^a{}_b(X) := \theta^a(\nabla_X e_b) \quad \Longleftrightarrow \quad \nabla_X e_b = \Gamma^a{}_b(X) e_a \;, \tag{A.9.9a}$$

$$\boxed{\Gamma^a{}_{bc} := \Gamma^a{}_b(e_c) = \theta^a(\nabla_{e_c} e_b)} \quad \Longleftrightarrow \quad \nabla_X e_b = \Gamma^a{}_{bc} X^c e_a \;. \tag{A.9.9b}$$

The (locally defined) functions $\Gamma^a{}_{bc}$ are called *connection coefficients*. If $\{e_a\}$ is the coordinate basis $\{\partial_\mu\}$ we shall write

$$\Gamma^\mu{}_{\alpha\beta} := dx^\mu(\nabla_{\partial_\beta}\partial_\alpha) \quad \left(\Longleftrightarrow \quad \nabla_{\partial_\mu}\partial_\nu = \Gamma^\sigma{}_{\nu\mu}\partial_\sigma \right) \;, \tag{A.9.10}$$

[1]This is the case when Ω is a coordinate patch with coordinates (x^i), then the $\{e_a\}_{a=1,\ldots,\dim M}$ can be chosen to be equal to $\{\partial_i\}_{a=1,\ldots,\dim M}$. Recall that a manifold is said to be parallelizable if a basis of TM can be chosen globally over M — in such a case Ω can be taken equal to M. We emphasize that we are *not* assuming that M is parallelizable, so that equations such as (A.9.9) have only a local character in general.

etc. In this particular case the connection coefficients are usually called *Christoffel symbols.* We will sometimes write $\Gamma^\sigma_{\nu\mu}$ instead of $\Gamma^\sigma{}_{\nu\mu}$; note that the former convention is more common. By using the Leibniz rule (A.9.5) we find

$$\begin{aligned}\nabla_X Y &= \nabla_X (Y^a e_a) \\ &= X(Y^a) e_a + Y^a \nabla_X e_a \\ &= X(Y^a) e_a + Y^a \Gamma^b{}_a(X) e_b \\ &= (X(Y^a) + \Gamma^a{}_b(X) Y^b) e_a \\ &= (X(Y^a) + \Gamma^a{}_{bc} Y^b X^c) e_a \ , \end{aligned} \tag{A.9.11}$$

which gives various equivalent ways of writing $\nabla_X Y$. The (perhaps only locally defined) $\Gamma^a{}_b$'s are linear in X, and the collection $(\Gamma^a{}_b)_{a,b=1,\ldots,\dim M}$ is sometimes referred to as the *connection one-form.* The one-covariant, one-contravariant tensor field ∇Y is defined as

$$\nabla Y := \nabla_a Y^b \theta^a \otimes e_b \Longleftrightarrow \nabla_a Y^b := \theta^b(\nabla_{e_a} Y) \Longleftrightarrow \boxed{\nabla_a Y^b = e_a(Y^b) + \Gamma^b{}_{ca} Y^c} \ . \tag{A.9.12}$$

We will often write ∇_a for ∇_{e_a}. Further, $\nabla_a Y^b$ will sometimes be written as $Y^b{}_{;a}$.

A.9.3 Transformation law

Consider a coordinate basis ∂_{x^i}, it is natural to enquire about the transformation law of the connection coefficients $\Gamma^i{}_{jk}$ under a change of coordinates $x^i \to y^k(x^i)$. To make things clear, let us write $\Gamma^i{}_{jk}$ for the connection coefficients in the x–coordinates, and $\hat\Gamma^i{}_{jk}$ for the ones in the y–cordinates. We calculate:

$$\begin{aligned}\Gamma^i{}_{jk} &:= dx^i\Big(\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j}\Big) \\ &= dx^i\Big(\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial y^\ell}{\partial x^j}\frac{\partial}{\partial y^\ell}\Big) \\ &= dx^i\Big(\frac{\partial^2 y^\ell}{\partial x^k \partial x^j}\frac{\partial}{\partial y^\ell} + \frac{\partial y^\ell}{\partial x^j}\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial y^\ell}\Big) \\ &= \frac{\partial x^i}{\partial y^s} dy^s\Big(\frac{\partial^2 y^\ell}{\partial x^k \partial x^j}\frac{\partial}{\partial y^\ell} + \frac{\partial y^\ell}{\partial x^j}\nabla_{\frac{\partial y^r}{\partial x^k}\frac{\partial}{\partial y^r}}\frac{\partial}{\partial y^\ell}\Big) \\ &= \frac{\partial x^i}{\partial y^s} dy^s\Big(\frac{\partial^2 y^\ell}{\partial x^k \partial x^j}\frac{\partial}{\partial y^\ell} + \frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^k}\nabla_{\frac{\partial}{\partial y^r}}\frac{\partial}{\partial y^\ell}\Big) \\ &= \frac{\partial x^i}{\partial y^s}\frac{\partial^2 y^s}{\partial x^k \partial x^j} + \frac{\partial x^i}{\partial y^s}\frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^k}\hat\Gamma^s{}_{\ell r} \ . \end{aligned} \tag{A.9.13}$$

Summarising,

$$\boxed{\Gamma^i{}_{jk} = \hat\Gamma^s{}_{\ell r}\frac{\partial x^i}{\partial y^s}\frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^k} + \frac{\partial x^i}{\partial y^s}\frac{\partial^2 y^s}{\partial x^k \partial x^j}} \ . \tag{A.9.14}$$

Thus, the $\Gamma^i{}_{jk}$'s do *not* form a tensor; instead they transform as a tensor *plus* a non-homogeneous term containing second derivatives, as seen above. However, because the inhomogeneous term in (A.9.14) is symmetric under the interchange of i and j, it follows from (A.9.14) that

$$T^i_{jk} := \Gamma^i{}_{kj} - \Gamma^i{}_{jk}$$

does transform as a tensor, called *the torsion tensor* of ∇.

EXERCICE A.9.1 Let $\Gamma^i{}_{jk}$ transform as in (A.9.14) under coordinate transformations. If X and Y are vector fields, define in local coordinates

$$\nabla_X Y := \Big(X(Y^i) + \Gamma^i{}_{jk} X^k Y^k\Big)\,\partial_i\ . \tag{A.9.15}$$

Show that $\nabla_X Y$ transforms as a vector field under coordinate transformations (and thus is a vector field).

A.9.4 Torsion

An index-free definition of torsion proceeds as follows: Let ∇ be a covariant derivative defined for vector fields, the *torsion tensor* T is defined by the formula

$$\boxed{T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]}\ , \tag{A.9.16}$$

where $[X,Y]$ is the Lie bracketWe obviously have

$$T(X,Y) = -T(Y,X)\ . \tag{A.9.17}$$

Let us check that T is actually a tensor field: multi-linearity with respect to addition is obvious. To check what happens under multiplication by functions, in view of (A.9.17) it is sufficient to do the calculation for the first slot of T. We then have

$$\begin{aligned} T(fX,Y) &= \nabla_{fX} Y - \nabla_Y(fX) - [fX,Y] \\ &= f\Big(\nabla_X Y - \nabla_Y X\Big) - Y(f)X - [fX,Y]\ . \end{aligned} \tag{A.9.18}$$

To work out the last commutator term we compute, for any function g,

$$[fX,Y](g) = fX(Y(g)) - \underbrace{Y(fX(g))}_{=Y(f)X(g)+fY(X(g))} = f[X,Y](g) - Y(f)X(g)\ ,$$

hence

$$[fX,Y] = f[X,Y] - Y(f)X\ , \tag{A.9.19}$$

and the last term here cancels the undesirable second-to-last term in (A.9.18), as required.

In a coordinate basis ∂_μ we have $[\partial_\mu, \partial_\nu] = 0$ and one finds from (A.9.10)

$$\boxed{T_{\mu\nu} := T(\partial_\mu, \partial_\nu) = (\Gamma^\sigma{}_{\nu\mu} - \Gamma^\sigma{}_{\mu\nu})\partial_\sigma}\ , \tag{A.9.20}$$

which shows that — in coordinate frames — T is determined by twice the antisymmetrization of the $\Gamma^\sigma{}_{\mu\nu}$'s over the lower indices. In particular that last antisymmetrization produces a tensor field.

A.9.5 Covectors

Suppose that we are given a covariant derivative on vector fields, there is a natural way of inducing a covariant derivative on one-forms by imposing the condition that *the duality operation be compatible with the Leibniz rule*: given two vector fields X and Y together with a field of one-forms α, one sets

$$\boxed{(\nabla_X\alpha)(Y) := X(\alpha(Y)) - \alpha(\nabla_X Y)} \,. \tag{A.9.21}$$

Let us, first, check that (A.9.21) indeed defines a field of one-forms. The linearity, in the Y variable, with respect to addition is obvious. Next, for any function f we have

$$\begin{aligned}(\nabla_X\alpha)(fY) &= X(\alpha(fY)) - \alpha(\nabla_X(fY)) \\ &= X(f)\alpha(Y) + fX(\alpha(Y)) - \alpha(X(f)Y + f\nabla_X Y) \\ &= f(\nabla_X\alpha)(Y)\,,\end{aligned}$$

as should be the case for one-forms. Next, we need to check that ∇ defined by (A.9.21) does satisfy the remaining axioms imposed on covariant derivatives. Again multi-linearity with respect to addition is obvious, as well as linearity with respect to multiplication of X by a function. Finally,

$$\begin{aligned}\nabla_X(f\alpha)(Y) &= X(f\alpha(Y)) - f\alpha(\nabla_X Y) \\ &= X(f)\alpha(Y) + f(\nabla_X\alpha)(Y)\,,\end{aligned}$$

as desired.

The duality pairing

$$T_p^*M \times T_pM \ni (\alpha, X) \to \alpha(X) \in \mathbb{R}$$

is sometimes called *contraction*. As already pointed out, the operation ∇ on one forms has been defined in (A.9.21) so as to satisfy the *Leibniz rule under duality pairing*:

$$X(\alpha(Y)) = (\nabla_X\alpha)(Y) + \alpha(\nabla_X Y)\,; \tag{A.9.22}$$

this follows directly from (A.9.21). This should not be confused with the Leibniz rule under multiplication by functions, which is part of the definition of a covariant derivative, and therefore always holds. It should be kept in mind that (A.9.22) does not necessarily hold for all covariant derivatives: if ${}^v\nabla$ is some covariant derivative on vectors, and ${}^f\nabla$ is some covariant derivative on one-forms, in general one will have

$$X(\alpha(Y)) \neq ({}^f\nabla_X)\alpha(Y) + \alpha({}^v\nabla_X Y)\,.$$

Using the basis-expression (A.9.11) of $\nabla_X Y$ and the definition (A.9.21) we have

$$\nabla_X\alpha = X^a\nabla_a\alpha_b\,\theta^b\,,$$

with

$$\begin{aligned}\boxed{\nabla_a\alpha_b} &:= (\nabla_{e_a}\alpha)(e_b) \\ &= e_a(\alpha(e_b)) - \alpha(\nabla_{e_a}e_b) \\ &= \boxed{e_a(\alpha_b) - \Gamma^c{}_{ba}\alpha_c}\,.\end{aligned}$$

A.9.6 Higher order tensors

It should now be clear how to extend ∇ to tensors of arbitrary valence: if T is r covariant and s contravariant one sets

$$\begin{aligned}(\nabla_X T)(X_1,\ldots,X_r,\alpha_1,\ldots\alpha_s) := X\Big(T(X_1,\ldots,X_r,\alpha_1,\ldots\alpha_s)\Big)\\ -T(\nabla_X X_1,\ldots,X_r,\alpha_1,\ldots\alpha_s) - \ldots - T(X_1,\ldots,\nabla_X X_r,\alpha_1,\ldots\alpha_s)\\ -T(X_1,\ldots,X_r,\nabla_X\alpha_1,\ldots\alpha_s) - \ldots - T(X_1,\ldots,X_r,\alpha_1,\ldots\nabla_X\alpha_s)\,.\end{aligned} \tag{A.9.23}$$

The verification that this defines a covariant derivative proceeds in a way identical to that for one-forms. In a basis we have

$$\nabla_X T = X^a \nabla_a T_{a_1\ldots a_r}{}^{b_1\ldots b_s}\theta^{a_1}\otimes\ldots\otimes\theta^{a_r}\otimes e_{b_1}\otimes\ldots\otimes e_{b_s}\ ,$$

and (A.9.23) gives

$$\begin{aligned}\nabla_a T_{a_1\ldots a_r}{}^{b_1\ldots b_s} &:= (\nabla_{e_a}T)(e_{a_1},\ldots,e_{a_r},\theta^{b_1},\ldots,\theta^{b_s})\\ &= e_a(T_{a_1\ldots a_r}{}^{b_1\ldots b_s}) - \Gamma^c{}_{a_1 a}T_{c\ldots a_r}{}^{b_1\ldots b_s} - \ldots - \Gamma^c{}_{a_r a}T_{a_1\ldots c}{}^{b_1\ldots b_s}\\ &\quad +\Gamma^{b_1}{}_{ca}T_{a_1\ldots a_r}{}^{c\ldots b_s} + \ldots + \Gamma^{b_s}{}_{ca}T_{a_1\ldots a_r}{}^{b_1\ldots c}\,.\end{aligned} \tag{A.9.24}$$

Carrying over the last two lines of (A.9.23) to the left-hand-side of that equation one obtains the Leibniz rule for ∇ under pairings of tensors with vectors or forms. It should be clear from (A.9.23) that ∇ defined by that equation is the *only covariant derivative which agrees with the original one on vectors, and which satisfies the Leibniz rule under the pairing operation.* We will only consider such covariant derivatives in this work.

A.9.7 Geodesics and Christoffel symbols

A geodesic can be defined as the stationary point of the action

$$I(\gamma) \quad=\quad \frac{1}{2}\int_a^b \underbrace{g(\dot\gamma,\dot\gamma)(s)}_{=:\mathscr{L}(\gamma,\dot\gamma)}\,ds\ , \tag{A.9.25}$$

where $\gamma:[a,b]\to M$ is a differentiable curve. Thus,

$$\mathscr{L}(x^\mu,\dot x^\nu) = \frac{1}{2}g_{\alpha\beta}(x^\mu)\dot x^\alpha\dot x^\beta\ .$$

One readily finds the Euler-Lagrange equations for this Lagrange function:

$$\frac{d}{ds}\left(\frac{\partial\mathscr{L}}{\partial\dot x^\mu}\right) = \frac{\partial\mathscr{L}}{\partial x^\mu} \quad\Longleftrightarrow\quad \frac{d^2x^\mu}{ds^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0\,. \tag{A.9.26}$$

This provides a very convenient way of calculating the Christoffel symbols: given a metric g, write down $\mathscr{L}$, work out the Euler-Lagrange equations, and identify the Christoffels as the coefficients of the first derivative terms in those equations.

(The Euler-Lagrange equations for (A.9.25) are identical with those of

$$\tilde{I}(\gamma) = \int_a^b \sqrt{|g(\dot\gamma,\dot\gamma)(s)|}ds \;, \tag{A.9.27}$$

but (A.9.25) is more convenient to work with. For example, $\mathscr{L}$ is differentiable at points where $\dot\gamma$ vanishes, while $\sqrt{|g(\dot\gamma,\dot\gamma)(s)|}$ is not. The aesthetic advantage of (A.9.27), of being reparameterization-invariant, is more than compensated by the calculational convenience of $\mathscr{L}$.)

Example A.9.2 As an example, consider a metric of the form

$$g = dr^2 + f(r)d\varphi^2 \;.$$

Special cases of this metric include the Euclidean metric on $\mathbb{R}^2$ (then $f(r) = r^2$), and the canonical metric on a sphere (then $f(r) = \sin^2 r$, with r actually being the polar angle θ). The Lagrangian (A.9.27) is thus

$$L = \frac{1}{2}\left(\dot r^2 + f(r)\dot\varphi^2\right) \;.$$

The Euler-Lagrange equations read

$$\underbrace{\frac{\partial L}{\partial \varphi}}_{0} = \frac{d}{ds}\left(\frac{\partial L}{\partial \dot\varphi}\right) = \frac{d}{ds}(f(r)\dot\varphi) \;,$$

so that

$$0 = f\ddot\varphi + f'\dot r\dot\varphi = f\left(\ddot\varphi + \Gamma^\varphi_{\varphi\varphi}\dot\varphi^2 + 2\Gamma^\varphi_{r\varphi}\dot r\dot\varphi + \Gamma^\varphi_r\dot r^2\right) \quad \Longrightarrow \quad \Gamma^\varphi_{\varphi\varphi} = \Gamma^\varphi_{rr} = 0 \;, \quad \Gamma^\varphi_{r\varphi} = \frac{f'}{2f} \;.$$

Similarly

$$\underbrace{\frac{\partial L}{\partial r}}_{f'\dot\varphi^2/2} = \frac{d}{ds}\left(\frac{\partial L}{\partial \dot r}\right) = \ddot r \;,$$

so that

$$\Gamma^r_{r\varphi} = \Gamma^r_{rr} = 0 \;, \quad \Gamma^r_{\varphi\varphi} = -\frac{f'}{2} \;.$$

A.10 The Levi-Civita connection

One of the fundamental results in pseudo-Riemannian geometry is that of the existence of a torsion-free connection which preserves the metric:

Theorem A.10.1 *Let g be a two-covariant symmetric non-degenerate tensor field on a manifold M. Then there exists a unique connection ∇ such that*

1. $\nabla g = 0$,

2. *the torsion tensor T of ∇ vanishes.*

PROOF: Using the definition of $\nabla_i g_{jk}$ we have

$$0 = \nabla_i g_{jk} \equiv \partial_i g_{jk} - \Gamma^\ell{}_{ji} g_{\ell k} - \Gamma^\ell{}_{ki} g_{\ell j} \; ; \qquad \text{(A.10.1)}$$

here we have written $\Gamma^i{}_{jk}$ instead of $\Gamma^i{}_{jk}$, as is standard in the literature. We rewrite this equation making cyclic permutations of indices, and changing the overall sign:

$$0 = -\nabla_j g_{ki} \equiv -\partial_j g_{ki} + \Gamma^\ell{}_{kj} g_{\ell i} + \Gamma^\ell{}_{ij} g_{\ell k} \; .$$
$$0 = -\nabla_k g_{ij} \equiv -\partial_k g_{ij} + \Gamma^\ell{}_{ik} g_{\ell j} + \Gamma^\ell{}_{jk} g_{\ell i} \; .$$

Adding the three equations and using symmetry of Γ^k_{ji} in ij one obtains

$$0 = \partial_i g_{jk} - \partial_j g_{ki} - \partial_k g_{ij} + 2\Gamma^\ell{}_{jk} g_{\ell i} \; ,$$

Multiplying by g^{im} we obtain

$$\Gamma^m{}_{jk} = g^{mi}\Gamma^\ell_{jk} g_{\ell i} = \frac{1}{2} g^{mi}(\partial_i g_{jk} - \partial_j g_{ki} - \partial_k g_{ij}) \; . \qquad \text{(A.10.2)}$$

This proves uniqueness.

A straightforward, though somewhat lengthy, calculation shows that the $\Gamma^m{}_{jk}$'s defined by (A.10.2) satisfy the transformation law (A.9.14). Exercice A.9.1 shows that the formula (A.9.15) defines a torsion-free connection. It then remains to check that the insertion of $\Gamma^m{}_{jk}$, as given by (A.10.2), into the right-hand-side of (A.10.1), indeed gives zero, proving existence. □

Let us give a coordinate-free version of the above, which turns out to be much messier: Suppose, first, that a connection satisfying the above is given. By the Leibniz rule we then have for any vector fields X, Y and Z,

$$0 = (\nabla_X g)(Y, Z) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) \; . \qquad \text{(A.10.3)}$$

One then rewrites the same equation applying cyclic permutations to X, Y, and Z, with a minus sign for the last equation:

$$\begin{aligned} g(\nabla_X Y, Z) + g(Y, \nabla_X Z) &= X(g(Y, Z)) \; , \\ g(\nabla_Y Z, X) + g(Z, \nabla_Y X) &= Y(g(Z, X)) \; , \\ -g(\nabla_Z X, Y) - g(X, \nabla_Z Y) &= -Z(g(X, Y)) \; . \end{aligned} \qquad \text{(A.10.4)}$$

As the torsion tensor vanishes, the sum of the left-hand-sides of these equations can be manipulated as follows:

$$\begin{aligned} & g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + g(\nabla_Y Z, X) + g(Z, \nabla_Y X) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \\ &= g(\nabla_X Y + \nabla_Y X, Z) + g(Y, \nabla_X Z - \nabla_Z X) + g(X, \nabla_Y Z - \nabla_Z Y) \\ &= g(2\nabla_X Y - [X, Y], Z) + g(Y, [X, Z]) + g(X, [Y, Z]) \\ &= 2g(\nabla_X Y, Z) - g([X, Y], Z) + g(Y, [X, Z]) + g(X, [Y, Z]) \; . \end{aligned}$$

This shows that the sum of the three equations (A.10.4) can be rewritten as

$$\begin{aligned} 2g(\nabla_X Y, Z) &= g([X, Y], Z) - g(Y, [X, Z]) - g(X, [Y, Z]) \\ &\quad + X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) \; . \end{aligned} \qquad \text{(A.10.5)}$$

Since Z is arbitrary and g is non-degenerate, the left-hand-side of this equation determines the vector field $\nabla_X Y$ uniquely, and uniqueness of ∇ follows.

To prove existence, let $S(X,Y)(Z)$ be defined as one half of the right-hand-side of (A.10.5),

$$\begin{aligned} S(X,Y)(Z) &= \frac{1}{2}\Big(X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) \\ &\quad + g(Z,[X,Y]) - g(Y,[X,Z]) - g(X,[Y,Z])\Big) . \end{aligned} \tag{A.10.6}$$

Clearly S is linear with respect to addition in all fields involved. It is straightforward to check that it is linear with respect to multiplication of Z by a function, and since g is non-degenerate there exists a unique vector field $W(X,Y)$ such that

$$S(X,Y)(Z) = g(W(X,Y),Z) .$$

One readily checks that the assignment

$$(X,Y) \to W(X,Y)$$

satisfies all the requirements imposed on a covariant derivative $\nabla_X Y$. With some more work one checks that ∇_X so defined is torsion free, and metric compatible. □

Let us check that (A.10.5) reproduces (A.10.2): Consider (A.10.5) with $X = \partial_\gamma$, $Y = \partial_\beta$ and $Z = \partial_\sigma$,

$$\begin{aligned} 2g(\nabla_\gamma \partial_\beta, \partial_\sigma) &= 2g(\Gamma^\rho{}_{\beta\gamma}\partial_\rho, \partial_\sigma) \\ &= 2g_{\rho\sigma}\Gamma^\rho{}_{\beta\gamma} \\ &= \partial_\gamma g_{\beta\sigma} + \partial_\beta g_{\gamma\sigma} - \partial_\sigma g_{\beta\gamma} \end{aligned} \tag{A.10.7}$$

Multiplying this equation by $g^{\alpha\sigma}/2$ we then obtain

$$\boxed{\Gamma^\alpha{}_{\beta\gamma} = \tfrac{1}{2} g^{\alpha\sigma}\{\partial_\beta g_{\sigma\gamma} + \partial_\gamma g_{\sigma\beta} - \partial_\sigma g_{\beta\gamma}\}} . \tag{A.10.8}$$

A.11 "Local inertial coordinates"

PROPOSITION A.11.1 *1. Let g be a Lorentzian metric, for every $p \in M$ there exists a neighborhood thereof with a coordinate system such that $g_{\mu\nu} = \eta_{\mu\nu} =$* $\mathrm{diag}(1,-1,\cdots,-1)$ *at p.*

2. If g is differentiable, then the coordinates can be further chosen so that

$$\partial_\sigma g_{\alpha\beta} = 0 \tag{A.11.1}$$

at p.

The coordinates above will be referred to as *local inertial coordinates near* p.

REMARK A.11.2 An analogous result holds for any pseudo-Riemannian metric. Note that the "normal coordinates" satisfy the above. However, for metrics of finite differentiability, the introduction of normal coordinates leads to a loss of differentiability of the metric components, while the construction below preserves the order of differentiability.

PROOF: 1. Let y^μ be any coordinate system around p, shifting by a constant vector we can assume that p corresponds to $y^\mu = 0$. Let $e_a = e_a{}^\mu \partial/\partial y^\mu$ be any frame at p such that $g(e_a, e_b) = \eta_{ab}$ — such frames can be found by, *e.g.*, a Gram-Schmidt orthogonalisation. Calculating the determinant of both sides of the equation

$$g_{\mu\nu} e_a{}^\mu e_b{}^\nu = \eta_{ab}$$

we obtain, at p,

$$\det(g_{\mu\nu}) \det(e_a{}^\mu)^2 = -1 \;,$$

which shows that $\det(e_a{}^\mu)$ is non-vanishing. It follows that the formula

$$y^\mu = e^\mu{}_a x^a$$

defines a (linear) diffeomorphism. In the new coordinates we have, again at p,

$$g\Big(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\Big) = e^\mu{}_a e^\nu{}_b g\Big(\frac{\partial}{\partial y^\mu}, \frac{\partial}{\partial y^\nu}\Big) = \eta_{ab} \;. \tag{A.11.2}$$

2. We will use (A.9.14), which uses latin indices, so let us switch to that notation. Let x^i be the coordinates described in point 1., recall that p lies at the origin of those coordinates. The new coordinates $\hat x^j$ will be implicitly defined by the equations

$$x^i = \hat x^i + \frac{1}{2} A^i{}_{jk} \hat x^j \hat x^k \;,$$

where $A^i{}_{jk}$ is a set of constants, symmetric with respect to the interchange of j and k. Recall (A.9.14),

$$\hat\Gamma^i{}_{jk} = \Gamma^s{}_{\ell r} \frac{\partial \hat x^i}{\partial x^s} \frac{\partial x^\ell}{\partial \hat x^j} \frac{\partial x^r}{\partial \hat x^k} + \frac{\partial \hat x^i}{\partial x^s} \frac{\partial^2 x^s}{\partial \hat x^k \partial \hat x^j} \;; \tag{A.11.3}$$

here we use $\hat\Gamma^s_{\ell r}$ to denote the Christoffel symbols of the metric in the hatted coordinates. Then, at $x^i = 0$, this equation reads

$$\begin{aligned} \hat\Gamma^i{}_{jk} &= \Gamma^s{}_{\ell r} \underbrace{\frac{\partial \hat x^i}{\partial x^s}}_{\delta^i_s} \underbrace{\frac{\partial x^\ell}{\partial \hat x^j}}_{\delta^\ell_j} \underbrace{\frac{\partial x^r}{\partial \hat x^k}}_{\delta^r_k} + \underbrace{\frac{\partial x^i}{\partial x^s}}_{\delta^i_s} \underbrace{\frac{\partial^2 x^s}{\partial \hat x^k \partial \hat x^j}}_{A^s_{kj}} \\ &= \Gamma^i{}_{jk} + A^i{}_{kj} \;. \end{aligned}$$

Choosing A^i_{jk} as $-\Gamma^i{}_{jk}(0)$, the result follows.

If you do not like to remember formulae such as (A.9.14), proceed as follows: Let x^μ be the coordinates described in point 1. The new coordinates $\hat x^\alpha$ will be implicitly defined by the equations

$$x^\mu = \hat x^\mu + \frac{1}{2} A^\mu{}_{\alpha\beta} \hat x^\alpha \hat x^\beta \;,$$

where $A^\mu{}_{\alpha\beta}$ is a set of constants, symmetric with respect to the interchange of α and β. Set

$$\hat g_{\alpha\beta} := g\Big(\frac{\partial}{\partial \hat x^\alpha}, \frac{\partial}{\partial \hat x^\beta}\Big) \;, \qquad g_{\alpha\beta} := g\Big(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\Big) \;.$$

Recall the transformation law

$$\hat{g}_{\mu\nu}(\hat{x}^\sigma) = g_{\alpha\beta}(x^\rho(\hat{x}^\sigma))\frac{\partial x^\alpha}{\partial \hat{x}^\mu}\frac{\partial x^\beta}{\partial \hat{x}^\nu} .$$

By differentiation one obtains at $x^\mu = \hat{x}^\mu = 0$,

$$\begin{aligned} \frac{\partial \hat{g}_{\mu\nu}}{\partial \hat{x}^\rho}(0) &= \frac{\partial g_{\mu\nu}}{\partial x^\rho}(0) + g_{\alpha\beta}(0)\Big(A^\alpha{}_{\mu\rho}\delta^\beta_\nu + \delta^\alpha_\mu A^\beta{}_{\nu\rho}\Big) \\ &= \frac{\partial g_{\mu\nu}}{\partial x^\rho}(0) + A_{\nu\mu\rho} + A_{\mu\nu\rho} , \end{aligned} \tag{A.11.4}$$

where

$$A_{\alpha\beta\gamma} := g_{\alpha\sigma}(0)A^\sigma{}_{\beta\gamma} .$$

It remains to show that we can choose $A^\sigma{}_{\beta\gamma}$ so that the left-hand-side can be made to vanish at p. An explicit formula for $A_{\sigma\beta\gamma}$ can be obtained from (A.11.4) by a cyclic permutation calculation similar to that in (A.10.4). After raising the first index, the final result is

$$A^\alpha{}_{\beta\gamma} = \frac{1}{2}g^{\alpha\rho}\left\{\frac{\partial g_{\beta\gamma}}{\partial x^\rho} - \frac{\partial g_{\beta\rho}}{\partial x^\gamma} - \frac{\partial g_{\rho\gamma}}{\partial x^\beta}\right\}(0) ;$$

the reader may wish to check directly that this does indeed lead to a vanishing right-hand-side of (A.11.4).

□

A.12 Curvature

Let ∇ be a covariant derivative defined for vector fields, the curvature tensor is defined by the formula

$$\boxed{R(X,Y)Z := \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z} , \tag{A.12.1}$$

where, as elsewhere, $[X,Y]$ is the Lie bracket defined in (A.3.6). We note the anti-symmetry

$$R(X,Y)Z = -R(Y,X)Z . \tag{A.12.2}$$

It turns out the this defines a tensor. Multi-linearity with respect to addition is obvious, but multiplication by functions require more work.

First, we have (see (A.9.19))

$$\begin{aligned} R(fX,Y)Z &= \nabla_{fX}\nabla_Y Z - \nabla_Y\nabla_{fX}Z - \nabla_{[fX,Y]}Z \\ &= f\nabla_X\nabla_Y Z - \nabla_Y(f\nabla_X Z) - \underbrace{\nabla_{f[X,Y]-Y(f)X}Z}_{=f\nabla_{[X,Y]}Z - Y(f)\nabla_X Z} \\ &= fR(X,Y)Z . \end{aligned}$$

The simplest proof of linearity in the last slot proceeds via an index calculation in adapted coordinates; so while we will do the "elegant", index-free version shortly, let us do the ugly one first. We use the coordinate system of Proposition A.11.1 below, in which the first derivatives of the metric vanish at the prescribed point p:

$$\begin{aligned}\nabla_i\nabla_j Z^k &= \partial_i(\partial_j Z^k - \Gamma^k{}_{\ell j}Z^\ell) + \underbrace{0\times\nabla Z}_{\text{at } p} \\ &= \partial_i\partial_j Z^k - \partial_i\Gamma^k{}_{\ell j}Z^\ell \quad \text{at } p\,. \end{aligned} \tag{A.12.3}$$

Antisymmetrising in i and j, the terms involving the second derivatives of Z drop out, so the result is indeed linear in Z. So $\nabla_i\nabla_j Z^k - \nabla_j\nabla_i Z^k$ is a tensor field linear in Z, and therefore can be written as $R^k{}_{\ell ij}Z^\ell$.

Note that $\nabla_i\nabla_j Z^k$ is, by definition, the tensor field of first covariant derivatives of the tensor field $\nabla_j Z^k$, while (A.12.1) involves covariant derivatives of vector fields only, so the equivalence of both approaches requires a further argument. This is provided in the calculation below leading to (A.12.5).

Next,

$$\begin{aligned} R(X,Y)(fZ) &= \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ) \\ &= \Big\{\nabla_X\Big(Y(f)Z + f\nabla_Y Z\Big)\Big\} - \Big\{\cdots\Big\}_{X\leftrightarrow Y} \\ &\quad -[X,Y](f)Z - f\nabla_{[X,Y]}Z \\ &= \Big\{\underbrace{X(Y(f))Z}_{a} + \underbrace{Y(f)\nabla_X Z + X(f)\nabla_Y Z}_{b} + f\nabla_X\nabla_Y Z\Big\} - \Big\{\cdots\Big\}_{X\leftrightarrow Y} \\ &\quad - \underbrace{[X,Y](f)Z}_{c} - f\nabla_{[X,Y]}Z\,. \end{aligned}$$

Now, a together with its counterpart with X and Y interchanged cancel out with c, while b is symmetric with respect to X and Y and therefore cancels out with its counterpart with X and Y interchanged, leading to the desired equality

$$R(X,Y)(fZ) \;=\; fR(X,Y)Z\,.$$

In a coordinate basis $\{e_a\} = \{\partial_\mu\}$ we find[2] (recall that $[\partial_\mu,\partial_\nu]=0$)

$$\begin{aligned} R^\alpha{}_{\beta\gamma\delta} &:= \langle dx^\alpha, R(\partial_\gamma,\partial_\delta)\partial_\beta\rangle \\ &= \langle dx^\alpha, \nabla_\gamma\nabla_\delta\partial_\beta\rangle - \langle\cdots\rangle_{\delta\leftrightarrow\gamma} \\ &= \langle dx^\alpha, \nabla_\gamma(\Gamma^\sigma{}_{\beta\delta}\partial_\sigma)\rangle - \langle\cdots\rangle_{\delta\leftrightarrow\gamma} \\ &= \langle dx^\alpha, \partial_\gamma(\Gamma^\sigma{}_{\beta\delta})\partial_\sigma + \Gamma^\rho{}_{\sigma\gamma}\Gamma^\sigma{}_{\beta\delta}\partial_\rho\rangle - \langle\cdots\rangle_{\delta\leftrightarrow\gamma} \\ &= \{\partial_\gamma\Gamma^\alpha{}_{\beta\delta} + \Gamma^\alpha{}_{\sigma\gamma}\Gamma^\sigma{}_{\beta\delta}\} - \{\cdots\}_{\delta\leftrightarrow\gamma}\,, \end{aligned}$$

leading finally to

$$\boxed{\boxed{R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma\Gamma^\alpha{}_{\beta\delta} - \partial_\delta\Gamma^\alpha{}_{\beta\gamma} + \Gamma^\alpha{}_{\sigma\gamma}\Gamma^\sigma{}_{\beta\delta} - \Gamma^\alpha{}_{\sigma\delta}\Gamma^\sigma{}_{\beta\gamma}}}\,. \tag{A.12.4}$$

In a general frame some supplementary commutator terms will appear in the formula for $R^a{}_{bcd}$.

We note the following:

[2] The reader is warned that certain authors use a different sign convention either for $R(X,Y)Z$, or for $R^\alpha{}_{\beta\gamma\delta}$, or both. A useful table that lists the sign conventions for a series of standard GR references can be found on the backside of the front cover of [125].

THEOREM A.12.1 *There exists a coordinate system in which the metric tensor field has vanishing second derivatives at p if and only if its Riemann tensor vanishes at p. Furthermore, there exists a coordinate system in which the metric tensor field has constant entries near p if and only if the Riemann tensor vanishes near p.*

PROOF: The condition is necessary, since Riem is a tensor. The sufficiency will be admitted. □

The calculation of the curvature tensor is often a very traumatic experience. There is one obvious case where things are painless, when all $g_{\mu\nu}$'s are constants: in this case the Christoffels vanish, and so does the curvature tensor.

For more general metrics one way out is to use symbolic computer algebra, this can, e.g., be done online on `http://grtensor.phy.queensu.ca/NewDemo`. The MATHEMATICA package xACT [114] provides a very powerful tool for all kinds of calculations involving curvature.

EXAMPLE A.12.2 As a less trivial example, consider the round two sphere, which we write in the form

$$g = d\theta^2 + e^{2f} d\varphi^2 \ , \qquad e^{2f} = \sin^2\theta \ .$$

The Christoffel symbols are easily founds from the Lagrangean for geodesics:

$$\mathscr{L} = \frac{1}{2}(\dot\theta^2 + e^{2f}\dot\varphi^2) \ .$$

The Euler-Lagrange equations give

$$\Gamma^\theta{}_{\varphi\varphi} = -f' e^{2f} \ , \quad \Gamma^\varphi{}_{\theta\varphi} = \Gamma^\varphi{}_{\varphi\theta} = f' \ ,$$

with the remaining Christoffel symbols vanishing. Using the definition of the Riemann tensor we then immediately find

$$R^\varphi{}_{\theta\varphi\theta} = -f'' - (f')^2 = 1 \ .$$

All remaining components of the Riemann tensor can be obtained from this one by raising and lowering of indices, together with the symmetry operations which we are about to describe. This leads to

$$R_{ab} = g_{ab} \ , \qquad R = 2 \ .$$

Equation (A.12.1) is most frequently used "upside-down", not as a definition of the Riemann tensor, but as a tool for calculating what happens when one changes the order of covariant derivatives. Recall that for partial derivatives we have

$$\partial_\mu\partial_\nu Z^\sigma = \partial_\nu\partial_\mu Z^\sigma \ ,$$

but this is not true in general if partial derivatives are replaced by covariant ones:

$$\nabla_\mu\nabla_\nu Z^\sigma \neq \nabla_\nu\nabla_\mu Z^\sigma \ .$$

To find the correct formula let us consider the tensor field S defined as

$$Y \longrightarrow S(Y) := \nabla_Y Z \ .$$

In local coordinates, S takes the form

$$S = \nabla_\mu Z^\nu \, dx^\mu \otimes \partial_\nu \; .$$

It follows from the Leibniz rule — or, equivalently, from the definitions in Section A.9 — that we have

$$\begin{aligned}(\nabla_X S)(Y) &= \nabla_X(S(Y)) - S(\nabla_X Y) \\ &= \nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z \; .\end{aligned}$$

The commutator of the derivatives can then be calculated as

$$\begin{aligned}(\nabla_X S)(Y) - (\nabla_Y S)(X) &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{\nabla_X Y} Z + \nabla_{\nabla_Y X} Z \\ &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ &\quad + \nabla_{[X,Y]} Z - \nabla_{\nabla_X Y} Z + \nabla_{\nabla_Y X} Z \\ &= R(X,Y)Z - \nabla_{T(X,Y)} Z \; . \end{aligned} \tag{A.12.5}$$

Writing ∇S in the usual form

$$\nabla S = \nabla_\sigma S_\mu{}^\nu \, dx^\sigma \otimes dx^\mu \otimes \partial_\nu = \nabla_\sigma \nabla_\mu Z^\nu \, dx^\sigma \otimes dx^\mu \otimes \partial_\nu \; ,$$

we are thus led to

$$\nabla_\mu \nabla_\nu Z^\alpha - \nabla_\nu \nabla_\mu Z^\alpha = R^\alpha{}_{\sigma\mu\nu} Z^\sigma - T^\sigma{}_{\mu\nu} \nabla_\sigma Z^\alpha \; . \tag{A.12.6}$$

In the important case of vanishing torsion, the coordinate-component equivalent of (A.12.1) is thus

$$\boxed{\nabla_\mu \nabla_\nu X^\alpha - \nabla_\nu \nabla_\mu X^\alpha = R^\alpha{}_{\sigma\mu\nu} X^\sigma} \; . \tag{A.12.7}$$

An identical calculation gives, still for torsionless connections,

$$\nabla_\mu \nabla_\nu a_\alpha - \nabla_\nu \nabla_\mu a_\alpha = -R^\sigma{}_{\alpha\mu\nu} a_\sigma \; . \tag{A.12.8}$$

For a general tensor t and torsion-free connection each tensor index comes with a corresponding Riemann tensor term:

$$\begin{aligned}&\nabla_\mu \nabla_\nu t_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\beta_s} - \nabla_\nu \nabla_\mu t_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\beta_s} = \\ &\quad -R^\sigma{}_{\alpha_1\mu\nu} t_{\sigma\ldots\alpha_r}{}^{\beta_1\ldots\beta_s} - \ldots - R^\sigma{}_{\alpha_r\mu\nu} t_{\alpha_1\ldots\sigma}{}^{\beta_1\ldots\beta_s} \\ &\quad + R^{\beta_1}{}_{\sigma\mu\nu} t_{\alpha_1\ldots\alpha_r}{}^{\sigma\ldots\beta_s} + \ldots + R^{\beta_s}{}_{\sigma\mu\nu} t_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\sigma} \; .\end{aligned} \tag{A.12.9}$$

A.12.1 Bianchi identities

We have already seen the anti-symmetry property of the Riemann tensor, which in the index notation corresponds to the equation

$$R^\alpha{}_{\beta\gamma\delta} = -R^\alpha{}_{\beta\delta\gamma} \; . \tag{A.12.10}$$

There are a few other identities satisfied by the Riemann tensor, we start with the *first Bianchi identity*. Let $A(X,Y,Z)$ be any expression depending upon three vector fields X,Y,Z which is antisymmetric in X and Y, we set

$$\sum_{[XYZ]} A(X,Y,Z) := A(X,Y,Z) + A(Y,Z,X) + A(Z,X,Y) \,, \qquad \text{(A.12.11)}$$

thus $\sum_{[XYZ]}$ is a sum over cyclic permutations of the vectors X,Y,Z. Clearly,

$$\sum_{[XYZ]} A(X,Y,Z) = \sum_{[XYZ]} A(Y,Z,X) = \sum_{[XYZ]} A(Z,X,Y) \,. \qquad \text{(A.12.12)}$$

Suppose, first, that X, Y and Z commute. Using (A.12.12) together with the definition (A.9.16) of the torsion tensor T we calculate

$$\begin{aligned}
\sum_{[XYZ]} R(X,Y)Z &= \sum_{[XYZ]} \Big(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z\Big) \\
&= \sum_{[XYZ]} \Big(\nabla_X\nabla_Y Z - \nabla_Y \underbrace{(\nabla_Z X + T(X,Z))}_{\text{we have used } [X,Z]=0, \text{ see (A.9.16)}}\Big) \\
&= \underbrace{\sum_{[XYZ]} \nabla_X\nabla_Y Z - \sum_{[XYZ]} \nabla_Y\nabla_Z X}_{=0 \text{ (see (A.12.12))}} - \sum_{[XYZ]} \nabla_Y(\underbrace{T(X,Z)}_{=-T(Z,X)}) \\
&= \sum_{[XYZ]} \nabla_X(T(Y,Z)) \,,
\end{aligned}$$

and in the last step we have again used (A.12.12). This can be somewhat rearranged by using the definition of the covariant derivative of a higher order tensor (compare (A.9.23)) — equivalently, using the Leibniz rule rewritten upside-down:

$$(\nabla_X T)(Y,Z) = \nabla_X(T(Y,Z)) - T(\nabla_X Y, Z) - T(Y, \nabla_X Z) \,.$$

This leads to

$$\begin{aligned}
\sum_{[XYZ]} \nabla_X(T(Y,Z)) &= \sum_{[XYZ]} \Big((\nabla_X T)(Y,Z) + T(\nabla_X Y, Z) + T(Y, \underbrace{\nabla_X Z}_{=T(X,Z)+\nabla_Z X})\Big) \\
&= \sum_{[XYZ]} \Big((\nabla_X T)(Y,Z) - T(\underbrace{T(X,Z)}_{=-T(Z,X)}, Y)\Big) \\
&\quad \underbrace{+ \sum_{[XYZ]} T(\nabla_X Y, Z) + \sum_{[XYZ]} \underbrace{T(Y, \nabla_Z X)}_{=-T(\nabla_Z X, Y)}}_{=0 \text{ (see (A.12.12))}} \\
&= \sum_{[XYZ]} \Big((\nabla_X T)(Y,Z) + T(T(X,Y),Z)\Big) \,.
\end{aligned}$$

Summarizing, we have obtained the first Bianchi identity:

$$\sum_{[XYZ]} R(X,Y)Z = \sum_{[XYZ]} \Big((\nabla_X T)(Y,Z) + T(T(X,Y),Z) \Big) \; , \qquad \text{(A.12.13)}$$

under the hypothesis that X, Y and Z commute. However, both sides of this equation are tensorial with respect to X, Y and Z, so that they remain correct without the commutation hypothesis.

We are mostly interested in connections with vanishing torsion, in which case (A.12.13) can be rewritten as

$$\boxed{R^\alpha{}_{\beta\gamma\delta} + R^\alpha{}_{\gamma\delta\beta} + R^\alpha{}_{\delta\beta\gamma} = 0} \; . \qquad \text{(A.12.14)}$$

Our next goal is the *second Bianchi identity*. We consider four vector fields X, Y, Z and W and we assume again that everybody commutes with everybody else. We calculate

$$\begin{aligned} \sum_{[XYZ]} \nabla_X (R(Y,Z)W) &= \sum_{[XYZ]} \Big(\underbrace{\nabla_X \nabla_Y \nabla_Z W}_{=R(X,Y)\nabla_Z W + \nabla_Y \nabla_X \nabla_Z W} - \nabla_X \nabla_Z \nabla_Y W \Big) \\ &= \sum_{[XYZ]} R(X,Y)\nabla_Z W \\ &\quad \underbrace{+ \sum_{[XYZ]} \nabla_Y \nabla_X \nabla_Z W - \sum_{[XYZ]} \nabla_X \nabla_Z \nabla_Y W}_{=0} \\ &\quad . \end{aligned} \qquad \text{(A.12.15)}$$

Next,

$$\begin{aligned} \sum_{[XYZ]} (\nabla_X R)(Y,Z)W &= \sum_{[XYZ]} \Big(\nabla_X (R(Y,Z)W) - R(\nabla_X Y, Z)W \\ &\quad - R(Y, \underbrace{\nabla_X Z}_{=\nabla_Z X + T(X,Z)})W - R(Y,Z)\nabla_X W \Big) \\ &= \sum_{[XYZ]} \nabla_X (R(Y,Z)W) \\ &\quad \underbrace{- \sum_{[XYZ]} R(\nabla_X Y, Z)W - \sum_{[XYZ]} \underbrace{R(Y, \nabla_Z X)W}_{=-R(\nabla_Z X, Y)W}}_{=0} \\ &\quad - \sum_{[XYZ]} \Big(R(Y, T(X,Z))W + R(Y,Z)\nabla_X W \Big) \\ &= \sum_{[XYZ]} \Big(\nabla_X (R(Y,Z)W) - R(T(X,Y),Z)W - R(Y,Z)\nabla_X W \Big) \; . \end{aligned}$$

It follows now from (A.12.15) that the first term cancels out the third one, leading to

$$\sum_{[XYZ]} (\nabla_X R)(Y,Z)W = - \sum_{[XYZ]} R(T(X,Y),Z)W \; , \qquad \text{(A.12.16)}$$

which is the desired second Bianchi identity for commuting vector fields. As before, because both sides are multi-linear with respect to addition and multiplication by functions, the result remains valid for arbitrary vector fields.

For torsionless connections the components equivalent of (A.12.16) reads

$$\boxed{R^\alpha{}_{\mu\beta\gamma;\delta} + R^\alpha{}_{\mu\gamma\delta;\beta} + R^\alpha{}_{\mu\delta\beta;\gamma} = 0} \,. \tag{A.12.17}$$

A.12.2 Pair interchange symmetry

There is one more identity satisfied by the curvature tensor which is specific to the curvature tensor associated with the Levi-Civita connection, namely

$$g(X, R(Y,Z)W) = g(Y, R(X,W)Z) \,. \tag{A.12.18}$$

If one sets

$$\boxed{R_{abcd} := g_{ae} R^e{}_{bcd}} \,, \tag{A.12.19}$$

then (A.12.18) is equivalent to

$$\boxed{R_{abcd} = R_{cdab}} \,. \tag{A.12.20}$$

We will present two proofs of (A.12.18). The first is direct, but not very elegant. The second is prettier, but less insightful.

For the ugly proof, we suppose that the metric is twice-differentiable. By point 2. of Proposition A.11.1, in a neighborhood of any point $p \in M$ there exists a coordinate system in which the connection coefficients $\Gamma^\alpha{}_{\beta\gamma}$ vanish at p. Equation (A.12.4) evaluated at p therefore reads

$$\begin{aligned}
R^\alpha{}_{\beta\gamma\delta} &= \partial_\gamma \Gamma^\alpha{}_{\beta\delta} - \partial_\delta \Gamma^\alpha{}_{\beta\gamma} \\
&= \frac{1}{2}\Big\{ g^{\alpha\sigma}\partial_\gamma(\partial_\delta g_{\sigma\beta} + \partial_\beta g_{\sigma\delta} - \partial_\sigma g_{\beta\delta}) \\
&\quad - g^{\alpha\sigma}\partial_\delta(\partial_\gamma g_{\sigma\beta} + \partial_\beta g_{\sigma\gamma} - \partial_\sigma g_{\beta\gamma})\Big\} \\
&= \frac{1}{2} g^{\alpha\sigma}\Big\{ \partial_\gamma\partial_\beta g_{\sigma\delta} - \partial_\gamma\partial_\sigma g_{\beta\delta} - \partial_\delta\partial_\beta g_{\sigma\gamma} + \partial_\delta\partial_\sigma g_{\beta\gamma} \Big\} \,.
\end{aligned}$$

Equivalently,

$$R_{\sigma\beta\gamma\delta}(0) = \frac{1}{2}\Big\{ \partial_\gamma\partial_\beta g_{\sigma\delta} - \partial_\gamma\partial_\sigma g_{\beta\delta} - \partial_\delta\partial_\beta g_{\sigma\gamma} + \partial_\delta\partial_\sigma g_{\beta\gamma} \Big\}(0) \,. \tag{A.12.21}$$

This last expression is obviously symmetric under the exchange of $\sigma\beta$ with $\gamma\delta$, leading to (A.12.20).

The above calculation traces back the pair-interchange symmetry to the definition of the Levi-Civita connection in terms of the metric tensor. As already mentioned, there exists a more elegant proof, where the origin of the symmetry is perhaps somewhat less apparent, which proceeds as follows: We start by noting that

$$0 = \nabla_a \nabla_b g_{cd} - \nabla_b \nabla_a g_{cd} = -R^e{}_{cab} g_{ed} - R^e{}_{dab} g_{ce} \,, \tag{A.12.22}$$

leading to anti-symmetry in the first two indices:

$$R_{abcd} = -R_{bacd} \ .$$

Next, using the cyclic symmetry for a torsion-free connection, we have

$$\begin{aligned} R_{abcd} + R_{cabd} + R_{bcad} &= 0 \ , \\ R_{bcda} + R_{dbca} + R_{cdba} &= 0 \ , \\ R_{cdab} + R_{acdb} + R_{dacb} &= 0 \ , \\ R_{dabc} + R_{bdac} + R_{abdc} &= 0 \ . \end{aligned}$$

The desired equation (A.12.20) follows now by adding the first two and subtracting the last two equations, using (A.12.22).

It is natural to enquire about the number of independent components of a tensor with the symmetries of a metric Riemann tensor in dimension n, the calculation proceeds as follows: as R_{abcd} is symmetric under the exchange of ab with cd, and anti-symmetric in each of these pairs, we can view it as a symmetric map from the space of anti-symmetric tensor with two indices. Now, the space of anti-symmetric tensors is $N = n(n-1)/2$ dimensional, while the space of symmetric maps in dimension N is $N(N+1)/2$ dimensional, so we obtain at most $n(n-1)(n^2-n+2)/8$ free parameters. However, we need to take into account the cyclic identity:

$$R_{abcd} + R_{bcad} + R_{cabd} = 0 \ . \qquad \text{(A.12.23)}$$

If $a = b$ this reads

$$R_{aacd} + R_{acad} + R_{caad} = 0 \ ,$$

which has already been accounted for. Similarly if $a = d$ we obtain

$$R_{abca} + R_{bcaa} + R_{caba} = 0 \ ,$$

which holds in view of the previous identities. We conclude that the only new identities which could possibly arise are those where $abcd$ are all distinct. Clearly no expression involving three such components of the Riemann tensor can be obtained using the previous identities, so this is an independent constraint. In dimension four (A.12.23) provides thus four candidate equations for another constraint, labeled by d, but it is easily checked that they all coincide; this leads to 20 free parameters at each space point. The reader is encouraged to finish the counting in higher dimensions.

A.13 Geodesic deviation (Jacobi equation)

Suppose that we have a one parameter family of geodesics

$$\gamma(s,\lambda) \text{ (in local coordinates, } (\gamma^\alpha(s,\lambda))),$$

where s is the parameter along the geodesic, and λ is a parameter which distinguishes the geodesics. Set

$$Z(s,\lambda) := \frac{\partial \gamma(s,\lambda)}{\partial \lambda} \equiv \frac{\partial \gamma^\alpha(s,\lambda)}{\partial \lambda} \partial_\alpha \ ;$$

for each λ this defines a vector field Z along $\gamma(s,\lambda)$, which measures how nearby geodesics deviate from each other, since, to first order, using a Taylor expansion,

$$\gamma^\alpha(s,\lambda) = \gamma^\alpha(s,\lambda_0) + Z^\alpha(\lambda-\lambda_0) + O((\lambda-\lambda_0)^2)\ .$$

To measure how a vector field W changes along $s \mapsto \gamma(s,\lambda)$, one introduces the differential operator D/ds, defined as

$$\frac{DW^\mu}{ds} := \frac{\partial(W^\mu\circ\gamma)}{\partial s} + \Gamma^\mu{}_{\alpha\beta}\dot\gamma^\beta W^\alpha \qquad \text{(A.13.1)}$$
$$= \dot\gamma^\beta\frac{\partial W^\mu}{\partial x^\beta} + \Gamma^\mu{}_{\alpha\beta}\dot\gamma^\beta W^\alpha \qquad \text{(A.13.2)}$$
$$= \dot\gamma^\beta\nabla_\beta W^\mu\ . \qquad \text{(A.13.3)}$$

(It would perhaps be more logical to write $\frac{DW^\mu}{\partial s}$ in the current context, but people never do that.) The last two lines only make sense if W is defined in a whole neighbourhood of γ, but for the first it suffices that $W(s)$ be defined along $s\mapsto\gamma(s,\lambda)$. (One possible way of making sense of the last two lines is to extend W^μ to any smooth vector field defined in a neighorhood of $\gamma^\mu(s,\lambda)$, and note that the result is independent of the particular choice of extension because the equation involves only derivatives tangential to $s\mapsto\gamma^\mu(s,\lambda)$.)

Analogously one sets

$$\frac{DW^\mu}{d\lambda} := \frac{\partial(W^\mu\circ\gamma)}{\partial\lambda} + \Gamma^\mu{}_{\alpha\beta}\partial_\lambda\gamma^\beta W^\alpha \qquad \text{(A.13.4)}$$
$$= \partial_\lambda\gamma^\beta\frac{\partial W^\mu}{\partial x^\beta} + \Gamma^\mu{}_{\alpha\beta}\partial_\lambda\gamma^\beta W^\alpha \qquad \text{(A.13.5)}$$
$$= Z^\beta\nabla_\beta W^\mu\ . \qquad \text{(A.13.6)}$$

Note that since $s\to\gamma(s,\lambda)$ is a geodesic we have from (A.13.1) and (A.13.3)

$$\frac{D^2\gamma^\mu}{ds^2} := \frac{D\dot\gamma^\mu}{ds} = \frac{\partial^2\gamma^\mu}{\partial s^2} + \Gamma^\mu{}_{\alpha\beta}\dot\gamma^\beta\dot\gamma^\alpha = 0\ . \qquad \text{(A.13.7)}$$

(This is sometimes written as $\dot\gamma^\alpha\nabla_\alpha\dot\gamma^\mu = 0$, which is again an abuse of notation since typically we will only know $\dot\gamma^\mu$ as a function of s, and so there is no such thing as $\nabla_\alpha\dot\gamma^\mu$.) Furthermore,

$$\frac{DZ^\mu}{ds} \underbrace{=}_{\text{(A.13.1)}} \frac{\partial^2\gamma^\mu}{\partial s\partial\lambda} + \Gamma^\mu{}_{\alpha\beta}\dot\gamma^\beta\partial_\lambda\gamma^\alpha \underbrace{=}_{\text{(A.13.4)}} \frac{D\dot\gamma^\mu}{d\lambda}\ , \qquad \text{(A.13.8)}$$

(The abuse-of-notation derivation of the same formula proceeds as:

$$\nabla_{\dot\gamma}Z^\mu = \dot\gamma^\nu\nabla_\nu Z^\mu = \dot\gamma^\nu\nabla_\nu\partial_\lambda\gamma^\mu \underbrace{=}_{\text{(A.13.3)}} \frac{\partial^2\gamma^\mu}{\partial s\partial\lambda} + \Gamma^\mu{}_{\alpha\beta}\dot\gamma^\beta\partial_\lambda\gamma^\alpha \underbrace{=}_{\text{(A.13.6)}} Z^\beta\nabla_\beta\dot\gamma^\mu = \nabla_Z\dot\gamma^\mu\ , \qquad \text{(A.13.9)}$$

which can then be written as

$$\nabla_{\dot\gamma}Z = \nabla_Z\dot\gamma\ .) \qquad \text{(A.13.10)}$$

One can now repeat the calculation leading to (A.12.7) to obtain, for any vector field W defined along $\gamma^\mu(s,\lambda)$,

$$\frac{D}{ds}\frac{D}{d\lambda}W^\mu - \frac{D}{d\lambda}\frac{D}{ds}W^\mu = R_{\alpha\beta\delta}{}^\mu\dot\gamma^\alpha Z^\beta W^\delta \ . \tag{A.13.11}$$

If $W^\mu = \dot\gamma^\mu$ the second term at the left-hand-side is zero, and from $\frac{D}{d\lambda}\dot\gamma = \frac{D}{ds}Z$ we obtain

$$\frac{D^2 Z^\mu}{ds^2}(s) = R_{\alpha\beta\sigma}{}^\mu\dot\gamma^\alpha Z^\beta\dot\gamma^\sigma \ . \tag{A.13.12}$$

We have obtained an equation known as the *Jacobi equation*, or as the *geodesic deviation equation*; in index-free notation:

$$\boxed{\frac{D^2 Z}{ds^2} = R(\dot\gamma, Z)\dot\gamma} \ . \tag{A.13.13}$$

Solutions of (A.13.13) are called *Jacobi fields* along γ.

A.14 Exterior algebra

A preferred class of tensors is provided by those that are totally antisymmetric in all indices. Such k-covariant tensors are called *k forms.* They are of special interest because they can naturally be used for integration. Furthermore, on such tensors one can introduce a differentiation operation, called *exterior derivative*, that does not require a connection.

Let α_i, $i = 1,\dots,k$, be a collection of one-forms, the *exterior product* of the α_i's is a k-form defined as

$$(\alpha_1\wedge\cdots\wedge\alpha_k)(X_1,\dots,X_k) = \det\left(\alpha_i(X_j)\right) , \tag{A.14.1}$$

where $\det\left(\alpha_i(X_j)\right)$ denotes the determinant of the matrix obtained by applying all the α_i's to all the vectors X_j. For example

$$dx^a\wedge dx^b(X,Y) = X^aY^b - Y^aX^a \ .$$

Note that this equals $dx^a\otimes dx^b - dx^b\otimes dx^a$, which is twice the antisymmetrisation $dx^{[a}\otimes dx^{b]}$. More generally, if α is a totaly anti-symmetric tensor with coordinate coefficients $\alpha_{a_1\dots a_k}$, then

$$\begin{aligned}\alpha &= \alpha_{a_1\dots a_k}dx^{a_1}\otimes\cdots\otimes dx^{a_k}\\ &= \alpha_{a_1\dots a_k}dx^{[a_1}\otimes\cdots\otimes dx^{a_k]}\\ &= \frac{1}{k!}\alpha_{a_1\dots a_k}dx^{a_1}\wedge\cdots\wedge dx^{a_k}\\ &= \sum_{a_1<\cdots<a_k}\alpha_{a_1\dots a_k}dx^{a_1}\wedge\cdots\wedge dx^{a_k} \ .\end{aligned}$$

This formula makes clear the factorial coefficients needed to go from tensor components to the components in the $dx^{a_1}\wedge\cdots\wedge dx^{a_k}$ basis.

A.15 Null hyperplanes and hypersurfaces

One of the objects that occur in Lorentzian geometry and which posses rather disturbing properties are *null hyperplanes* and *null hypersurfaces*, and it appears useful to include a short discussion of those. Perhaps the most unusual feature of such objects is that the direction normal is actually tangential as well. Furthermore, because the normal has no natural normalization, there is no natural measure induced on a null hypersurface by the ambient metric.

We start with some algebraic preliminaries. Let W be a real vector space, and recall that its dual W^* is defined as the set of all linear maps from W to $\mathbb{R}$ in the applications (in this work only vector spaces over the reals are relevant, but the field makes no difference for the discussion below). To avoid unnecessary complications we assume that W is finite dimensional. It is then standard that W^* has the same dimension as W.

We suppose that W is equipped with a *a) bilinear, b) symmetric, and c) non-degenerate form* q. Thus

$$q : W \to W$$

satisfies

$$a) \quad q(\lambda X + \mu Y, Z) = \lambda q(X, Z) + \mu q(Y, Z) \ , \qquad b) \quad q(X, Y) = q(Y, X) \ ,$$

and we also have the implication

$$c) \quad \forall Y \in W \ q(X, Y) = 0 \implies X = 0 \ . \tag{A.15.1}$$

(Strictly speaking, we should have indicated linearity with respect to the second variable in a) as well, but this property follows from a) and b) as above). By an abuse of terminology, we will call q a *scalar product*; note that standard algebra textbooks often add the condition of positive-definiteness to the definition of scalar product, which we do not include here.

Let $V \subset W$ be a vector subspace of W. The *annihilator* V^0 of W is defined as the set of linear forms on W which vanish on V:

$$V^0 := \{\alpha \in W^* : \ \forall Y \in V \quad \alpha(Y) = 0\} \subset W^* \ .$$

V^0 is obviously a linear subspace of W^*.

Because q non-degenerate, it defines a linear isomorphism, denoted by $\flat$, between W and W^* by the formula:

$$X^\flat(Y) = g(X, Y) \ .$$

Indeed, the map $X \mapsto X^\flat$ is clearly linear. Next, it has no kernel by (A.15.1). Since the dimensions of W and W^* are the same, it must be an isomorphism. The inverse map is denoted by $\sharp$. Thus, by definition we have

$$g(\alpha^\sharp, Y) = \alpha(Y) \ .$$

The map $\flat$ is nothing but "the lowering of the index on a vector using the metric q", while $\sharp$ is the "raising of the index on a one-form using the inverse metric".

For further purposes it is useful to recall the standard fact:

PROPOSITION A.15.1

$$\dim V + \dim V^0 = \dim W \ .$$

PROOF: Let $\{e_i\}_{i=1,\dots,\dim V}$ be any basis of V, we can complete $\{e_i\}$ to a basis $\{e_i, f_a\}$, with $a = 1, \dots, \dim W - \dim V$, of W. Let $\{e_i^*, f_a^*\}$ be the dual basis of W^*. It is straightforward to check that V^0 is spanned by $\{f_a^*\}$, which gives the result. □

The quadratic form q defines the notion of orthogonality:

$$V^\perp := \{Y \in W : \forall X \in V \ g(X,Y) = 0\} \ .$$

A chase through the definitions above shows that

$$V^\perp = (V^0)^\sharp \ .$$

Proposition A.15.1 implies:

PROPOSITION A.15.2

$$\dim V + \dim V^\perp = \dim W \ .$$

This implies, again regardless of signature:

PROPOSITION A.15.3

$$(\dim V^\perp)^\perp = V \ .$$

PROOF: The inclusion $(\dim V^\perp)^\perp \supset V$ is obvious from the definitions. The equality follows now because both spaces have the same dimension, as a consequence of Proposition (A.15.2). □

Now,

$$X \in V \cap V^\perp \implies q(X,X) = 0 \ , \qquad \text{(A.15.2)}$$

so that X vanishes if q is positive- or negative-definite, leading to $\dim V \cap \dim V^\perp = \{0\}$ in those cases. However, this does not have to be the case anymore for non-definite scalar products q.

A vector subspace V of W is called a *hyperplane* if

$$\dim V = \dim W - 1 \ .$$

Proposition A.15.2 implies then

$$\dim V^\perp = 1 \ ,$$

regardless of the signature of q. Thus, given a hyperplane V there exists a vector w such that

$$V^\perp = \mathbb{R}w \ .$$

If q is Lorentzian, we say that

$$V \text{ is } \begin{cases} \text{spacelike} & \text{if } w \text{ is timelike;} \\ \text{timelike} & \text{if } w \text{ is spacelike;} \\ \text{null} & \text{if } w \text{ is null.} \end{cases}$$

An argument based e.g. on Gram-Schmidt orthonormalization shows that if V is spacelike, then the scalar product defined on V by restriction is positive-definite; similarly if V is timelike, then the resulting scalar product is Lorentzian. The last case, of a null V, leads to a degenerate induced scalar product. In fact, we claim that

$$V \text{ is null if and only if } V \text{ contains its normal.} . \tag{A.15.3}$$

To see (A.15.3), suppose that $V^\perp = \mathbb{R}w$, with w null. Since $g(w,w) = 0$ we have $w \in (\mathbb{R}w)^\perp$, and from Proposition A.15.3

$$w \in (\mathbb{R}w)^\perp = (V^\perp)^\perp = V \, .$$

Since V does not contain its normal in the remaining cases, the equivalence is established.

A hypersurface is $\mathscr{N} \subset \mathscr{M}$ called *null* if at every $p \in \mathscr{N}$ the tangent space $T_p\mathscr{N}$ is a null subspace of $T_p\mathscr{M}$. So (A.15.2) shows that a normal to a null hypersurface $\mathscr{N}$ is also tangent to $\mathscr{N}$.

A.16 Moving frames

A formalism which is very convenient for practical calculations is that of *moving frames*; it also plays a key role when considering spinors. By definition, a moving frame is a (locally defined) field of bases $\{e_a\}$ of TM such that the scalar products

$$g_{ab} := g(e_a, e_b) \tag{A.16.1}$$

are point independent. In most standard applications one assumes that the e_a's form an orthonormal basis, so that g_{ab} is a diagonal matrix with plus and minus ones on the diagonal. However, it is sometimes convenient to allow other such frames, *e.g.* with isotropic vectors being members of the frame.

It is customary to denote by ${\omega^a}_{bc}$ the associated connection coefficients:

$${\omega^a}_{bc} := \theta^a(\nabla_{e_c} e_b) \quad \Longleftrightarrow \quad \nabla_X e_b = {\omega^a}_{bc} X^c e_a \, , \tag{A.16.2}$$

where, as elsewhere, $\{\theta^a(p)\}$ is a basis of T^*_pM dual to $\{e_a(p)\} \subset T_pM$; we will refer to θ^a as a *coframe*. The *connection one forms* ${\omega^a}_b$ are defined as

$${\omega^a}_b(X) := \theta^a(\nabla_X e_b) \quad \Longleftrightarrow \quad \nabla_X e_b = {\omega^a}_b(X) e_a \, ; . \tag{A.16.3}$$

As always we use the metric to raise and lower indices, even though the ${\omega^a}_{bc}$'s do not form a tensor, so that

$$\omega_{abc} := g_{ad} {\omega^e}_{bc} \, , \quad \omega_{ab} := g_{ae} {\omega^e}_b \, . \tag{A.16.4}$$

When ∇ is metric compatible, the ω_{ab}'s are anti-antisymmetric: indeed, as the g_{ab}'s are point independent, for any vector field X we have

$$\begin{aligned}
0 = X(g_{ab}) = X(g(e_a, e_b)) &= g(\nabla_X e_a, e_b) + g(e_a, \nabla_X e_b) \\
&= g(\omega^c{}_a(X)e_c, e_b) + g(e_a, \omega^d{}_b(X)e_d) \\
&= g_{cb}\omega^c{}_a(X) + g_{ad}\omega^d{}_b(X) \\
&= \omega_{ba}(X) + \omega_{ab}(X) \ .
\end{aligned}$$

Hence

$$\boxed{\omega_{ab} = -\omega_{ba} \quad \Longleftrightarrow \quad \omega_{abc} = -\omega_{bac}} \ . \tag{A.16.5}$$

One can obtain a formula for the ω_{ab}'s in terms of Christoffels, the frame vectors and their derivatives: In order to see this, we note that

$$g(e_a, \nabla_{e_c} e_b) = g(e_a, \omega^d{}_{bc} e_d) = g_{ad}\omega^d{}_{bc} = \omega_{abc} \ . \tag{A.16.6}$$

Rewritten the other way round this gives an alternative equation for the ω's with all indices down:

$$\omega_{abc} = g(e_a, \nabla_{e_c} e_b) \quad \Longleftrightarrow \quad \omega_{ab}(X) = g(e_a, \nabla_X e_b) \ . \tag{A.16.7}$$

Then, writing

$$e_a = e_a{}^\mu \partial_\mu \ ,$$

we find

$$\begin{aligned}
\omega_{abc} &= g(e_a{}^\mu \partial_\mu, e_c{}^\lambda \nabla_\lambda e_b) \\
&= g_{\mu\sigma} e_a{}^\mu e_c{}^\lambda (\partial_\lambda e_b{}^\sigma + \Gamma^\sigma_{\lambda\nu} e_b{}^\nu) \ .
\end{aligned} \tag{A.16.8}$$

Next, it turns out that we can calculate the ω_{ab}'s in terms of the Lie brackets of the vector fields e_a, without having to calculate the Christoffel symbols. This shouldn't be too surprising, since an ON frame defines the metric uniquely. If ∇ has no torsion, from (A.16.7) we find

$$\omega_{abc} - \omega_{acb} = g(e_a, \nabla_{e_c} e_b - \nabla_{e_b} e_c) = g(e_a, [e_c, e_b]) \ .$$

We can now carry-out the usual cyclic permutations calculation to obtain

$$\begin{aligned}
\omega_{abc} - \omega_{acb} &= g(e_a, [e_c, e_b]) \ , \\
-(\omega_{bca} - \omega_{bac}) &= -g(e_b, [e_a, e_c]) \ , \\
-(\omega_{cab} - \omega_{cba}) &= -g(e_c, [e_b, e_a]) \ .
\end{aligned}$$

So, if the connection is the Levi-Civita connection, summing the three equations and using (A.16.5) leads to

$$\boxed{\omega_{abc} = \frac{1}{2}\Big(g(e_a, [e_c, e_b]) - g(e_b, [e_a, e_c]) - g(e_c, [e_b, e_a])\Big)} \ . \tag{A.16.9}$$

Equations (A.16.8)-(A.16.9) provide explicit expressions for the ω's. While it is useful to know that there are such expressions, and while those expressions are useful to estimate things for PDE purposes, they are rarely used for practical calculations; see Example A.16.2 for more comments about that last issue.

It turns out that one can obtain a simple expression for the torsion of ω using exterior differentiation. Recall that if α is a one-form, then its exterior derivative $d\alpha$ can be defined using the formula

$$d\alpha(X,Y) = X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X,Y]) \ . \tag{A.16.10}$$

We set

$$T^a(X,Y) := \theta^a(T(X,Y)) \ ,$$

and using (A.16.10) together with the definition (A.9.16) of the torsion tensor T we calculate as follows:

$$\begin{aligned} T^a(X,Y) &= \theta^a(\nabla_X Y - \nabla_Y X - [X,Y]) \\ &= X(Y^a) + \omega^a{}_b(X)Y^b - Y(X^a) - \omega^a{}_b(Y)X^b - \theta^a([X,Y]) \\ &= X(\theta^a(Y)) - Y(\theta^a(X)) - \theta^a([X,Y]) + \omega^a{}_b(X)\theta^b(Y) - \omega^a{}_b(Y)\theta^b(X) \\ &= d\theta^a(X,Y) + (\omega^a{}_b \wedge \theta^b)(X,Y) \ . \end{aligned}$$

It follows that

$$T^a = d\theta^a + \omega^a{}_b \wedge \theta^b \ . \tag{A.16.11}$$

In particular when the torsion vanishes we obtain the so-called *Cartan's first structure equation*

$$\boxed{d\theta^a + \omega^a{}_b \wedge \theta^b = 0} \ . \tag{A.16.12}$$

EXAMPLE A.16.1 As a simple example, we consider a two-dimensional metric of the form

$$g = dx^2 + e^{2f} dy^2 \ , \tag{A.16.13}$$

where f could possibly depend upon x and y. A natural frame is given by

$$\theta^1 = dx \ , \quad \theta^2 = e^f dy \ .$$

The first Cartan structure equations read

$$0 = \underbrace{d\theta^1}_{0} + \omega^1{}_b \wedge \theta^b = \omega^1{}_2 \wedge \theta^2 \ ,$$

since $\omega^1{}_1 = \omega_{11} = 0$ by antisymmetry, and

$$0 = \underbrace{d\theta^2}_{e^f \partial_x f dx \wedge dy} + \omega^1{}_b \wedge \theta^b = \partial_x f \theta^1 \wedge \theta^2 + \omega^2{}_1 \wedge \theta^1 \ .$$

It should then be clear that both equations can be solved by choosing ω_{12} proportional to θ^2, and such an ansatz leads to

$$\omega_{12} = -\omega_{21} = -\partial_x f \, \theta^2 = -\partial_x(e^f)\, dy \ . \tag{A.16.14}$$

EXAMPLE A.16.2 As another example of the moving frame technique we consider (the most general) three-dimensional spherically symmetric metric

$$g = e^{2\beta(r)}dr^2 + e^{2\gamma(r)}d\theta^2 + e^{2\gamma(r)}\sin^2\theta d\varphi^2\ . \tag{A.16.15}$$

There is an obvious choice of ON coframe for g given by

$$\theta^1 = e^{\beta(r)}dr\ ,\ \theta^2 = e^{\gamma(r)}d\theta\ ,\ \theta^3 = e^{\gamma(r)}\sin\theta d\varphi\ , \tag{A.16.16}$$

leading to

$$g = \theta^1 \otimes\ \theta^1 + \theta^2 \otimes\ \theta^2 + \theta^3 \otimes\ \theta^3\ ,$$

so that the frame e_a dual to the θ^a's will be ON, as desired:

$$g_{ab} = g(e_a, e_b) = \mathrm{diag}(1,1,1)\ .$$

The idea of the calculation which we are about to do is the following: there is only one connection which is compatible with the metric, and which is torsion free. If we find a set of one forms ω_{ab} which exhibit the properties just mentioned, then they have to be the connection forms of the Levi-Civita connection. As shown in the calculation leading to (A.16.5), the compatibility with the metric will be ensured if we require

$$\omega_{11} = \omega_{22} = \omega_{33} = 0\ ,$$

$$\omega_{12} = -\omega_{21}\ ,\quad \omega_{13} = -\omega_{31}\ ,\quad \omega_{23} = -\omega_{32}\ .$$

Next, we have the equations for the vanishing of torsion:

$$\begin{aligned}
0 = d\theta^1 &= -\underbrace{\omega^1{}_1}_{=0}\theta^1 - \omega^1{}_2\theta^2 - \omega^1{}_3\theta^3 \\
&= -\omega^1{}_2\theta^2 - \omega^1{}_3\theta^3\ , \\
d\theta^2 &= \gamma' e^\gamma dr \wedge d\theta = \gamma' e^{-\beta}\theta^1 \wedge \theta^2 \\
&= -\underbrace{\omega^2{}_1}_{=-\omega^1{}_2}\theta^1 - \underbrace{\omega^2{}_2}_{=0}\theta^2 - \omega^2{}_3\theta^3 \\
&= \omega^1{}_2\theta^1 - \omega^2{}_3\theta^3\ , \\
d\theta^3 &= \gamma' e^\gamma \sin\theta\, dr \wedge d\varphi + e^\gamma \cos\theta\, d\theta \wedge d\varphi = \gamma' e^{-\beta}\theta^1 \wedge \theta^3 + e^{-\gamma}\cot\theta\, \theta^2 \wedge \theta^3 \\
&= -\underbrace{\omega^3{}_1}_{=-\omega^1{}_3}\theta^1 - \underbrace{\omega^3{}_2}_{=-\omega^2{}_3}\theta^2 - \underbrace{\omega^3{}_3}_{=0}\theta^3 \\
&= \omega^1{}_3\theta^1 + \omega^2{}_3\theta^2\ .
\end{aligned}$$

Summarising,

$$\begin{aligned}
-\omega^1{}_2\theta^2 - \omega^1{}_3\theta^3 &= 0\ , \\
\omega^1{}_2\theta^1 \qquad - \omega^2{}_3\theta^3 &= \gamma' e^{-\beta}\theta^1 \wedge \theta^2\ , \\
\omega^1{}_3\theta^1 + \omega^2{}_3\theta^2 &= \gamma' e^{-\beta}\theta^1 \wedge \theta^3 + e^{-\gamma}\cot\theta\, \theta^2 \wedge \theta^3\ .
\end{aligned}$$

It should be clear from the first and second line that an $\omega^1{}_2$ proportional to θ^2 should do the job; similarly from the first and third line one sees that an $\omega^1{}_3$ proportional to θ^3 should work. It is then easy to find the relevant coefficient, as well as to find $\omega^2{}_3$:

$$\omega^1{}_2 = -\gamma' e^{-\beta}\theta^2 = -\gamma' e^{-\beta+\gamma}d\theta\ , \tag{A.16.17a}$$
$$\omega^1{}_3 = -\gamma' e^{-\beta}\theta^3 = -\gamma' e^{-\beta+\gamma}\sin\theta\, d\varphi\ , \tag{A.16.17b}$$
$$\omega^2{}_3 = -e^{-\gamma}\cot\theta\, \theta^3 = -\cos\theta\, d\varphi\ . \tag{A.16.17c}$$

It is convenient to define *curvature two-forms*:

$$\Omega^a{}_b = R^a{}_{bcd}\theta^c \otimes \theta^d = \frac{1}{2}R^a{}_{bcd}\theta^c \wedge \theta^d \ . \tag{A.16.18}$$

The *second Cartan structure equation* then reads

$$\boxed{\Omega^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b} \ . \tag{A.16.19}$$

This identity is easily verified using (A.16.10):

$$\begin{aligned}
\Omega^a{}_b(X,Y) &= \frac{1}{2}R^a{}_{bcd}\underbrace{\theta^c \wedge \theta^d(X,Y)}_{=X^cY^d - X^dY^c} \\
&= R^a{}_{bcd}X^cY^d \\
&= \theta^a(\nabla_X\nabla_Y e_b - \nabla_Y\nabla_X e_b - \nabla_{[X,Y]}e_b) \\
&= \theta^a(\nabla_X(\omega^c{}_b(Y)e_c) - \nabla_Y(\omega^c{}_b(X)e_c) - \omega^c{}_b([X,Y])e_c) \\
&= \theta^a\Big(X(\omega^c{}_b(Y))e_c + \omega^c{}_b(Y)\nabla_X e_c \\
&\quad -Y(\omega^c{}_b(X))e_c - \omega^c{}_b(X)\nabla_Y e_c - \omega^c{}_b([X,Y])e_c\Big) \\
&= X(\omega^a{}_b(Y)) + \omega^c{}_b(Y)\omega^a{}_c(X) \\
&\quad -Y(\omega^a{}_b(X)) - \omega^c{}_b(X)\omega^a{}_c(Y) - \omega^a{}_b([X,Y]) \\
&= \underbrace{X(\omega^a{}_b(Y)) - Y(\omega^a{}_b(X)) - \omega^a{}_b([X,Y])}_{=d\omega^a{}_b(X,Y)} \\
&\quad +\omega^a{}_c(X)\omega^c{}_b(Y) - \omega^a{}_c(Y)\omega^c{}_b(X) \\
&= (d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b)(X,Y) \ .
\end{aligned}$$

Equation (A.16.19) provides an efficient way of calculating the curvature tensor of any metric.

Example A.16.1 continued We have seen that the connection one-forms for the metric

$$g = dx^2 + e^{2f}dy^2 \tag{A.16.20}$$

read

$$\omega_{12} = -\omega_{21} = -\partial_x f\,\theta^2 = -\partial_x(e^f)\,dy \ .$$

By symmetry the only non-vanishing curvature two-forms are $\Omega_{12} = -\Omega_{21}$. From (A.16.19) we find

$$\Omega_{12} = d\omega_{12} + \underbrace{\omega_{1b} \wedge \omega^b{}_2}_{=\omega_{12}\wedge\omega^2{}_2=0} = -\partial_x^2(e^f)\,dx \wedge dy = -e^{-f}\partial_x^2(e^f)\,\theta^1 \wedge \theta^2 \ .$$

We conclude that

$$R_{1212} = -e^{-f}\partial_x^2(e^f) \ . \tag{A.16.21}$$

(Compare Example A.12.2, p. 92.) For instance, if g is the unit round metric on the two-sphere, then $f = \sin x$, and $R_{1212} = 1$. If $f = \sinh x$, then g is the canonical metric on hyperbolic space, and $R_{1212} = -1$. Finally, the function $f = \cosh x$ defines a *hyperbolic wormhole*, with again $R_{1212} = -1$.

EXAMPLE A.16.2 CONTINUED: From (A.16.17) we find:

$$
\begin{aligned}
\Omega^1{}_2 &= d\omega^1{}_2 + \underbrace{\omega^1{}_1}_{=0} \wedge \omega^1{}_2 + \omega^1{}_2 \wedge \underbrace{\omega^2{}_2}_{=0} + \underbrace{\omega^1{}_3 \wedge \omega^3{}_2}_{\sim \theta^3 \wedge \theta^3 = 0} \\
&= -d(\gamma' e^{-\beta+\gamma} d\theta) \\
&= -(\gamma' e^{-\beta+\gamma})' dr \wedge d\theta \\
&= -(\gamma' e^{-\beta+\gamma})' e^{-\beta-\gamma} \theta^1 \wedge \theta^2 \\
&= \sum_{a<b} R^1{}_{2ab} \theta^a \wedge \theta^b \ ,
\end{aligned}
$$

which shows that the only non-trivial coefficient (up to permutations) with the pair 12 in the first two slots is

$$R^1{}_{212} = -(\gamma' e^{-\beta+\gamma})' e^{-\beta-\gamma} \ . \tag{A.16.22}$$

A similar calculation, or arguing by symmetry, leads to

$$R^1{}_{313} = -(\gamma' e^{-\beta+\gamma})' e^{-\beta-\gamma} \ . \tag{A.16.23}$$

Finally,

$$
\begin{aligned}
\Omega^2{}_3 &= d\omega^2{}_3 + \omega^2{}_1 \wedge \omega^1{}_3 + \underbrace{\omega^2{}_2}_{=0} \wedge \omega^2{}_3 + \omega^2{}_3 \wedge \underbrace{\omega^3{}_3}_{=0} \\
&= -d(\cos\theta \, d\varphi) + (\gamma' e^{-\beta} \theta^2) \wedge (-\gamma' e^{-\beta} \theta^3) \\
&= (e^{-2\gamma} - (\gamma')^2 e^{-2\beta}) \theta^2 \wedge \theta^3 \ ,
\end{aligned}
$$

yielding

$$R^2{}_{323} = e^{-2\gamma} - (\gamma')^2 e^{-2\beta} \ . \tag{A.16.24}$$

The curvature scalar can easily be calculated now to be

$$
\begin{aligned}
R = R^{ij}{}_{ij} &= 2(R^{12}{}_{12} + R^{13}{}_{13} + R^{23}{}_{23}) \\
&= -4(\gamma' e^{-\beta+\gamma})' e^{-\beta-\gamma} + 2(e^{-2\gamma} - (\gamma')^2 e^{-2\beta}) \ . \qquad \text{(A.16.25)}
\end{aligned}
$$

Appendix B

Weyl connections, conformal rescalings of the metric

Consider a metric $\widetilde{g}$ related to g by a conformal rescaling:

$$\widetilde{g}_{ij} = \varphi^\ell g_{ij} \quad \Longleftrightarrow \quad \widetilde{g}^{ij} = \varphi^{-\ell} g^{ij} \ , \tag{B.0.1}$$

where φ is a function and ℓ is a real number. This gives the following transformation law for the Christoffel symbols:

$$\begin{aligned}
\widetilde{\Gamma}^i{}_{jk} &= \frac{1}{2}\widetilde{g}^{im}(\partial_j \widetilde{g}_{km} + \partial_k \widetilde{g}_{jm} - \partial_m \widetilde{g}_{jk}) \\
&= \frac{1}{2}\varphi^{-\ell} g^{im}(\partial_j(\varphi^\ell \widetilde{g}_{km}) + \partial_k(\varphi^\ell \widetilde{g}_{jm}) - \partial_m(\varphi^\ell g_{jk})) \\
&= \Gamma^i{}_{jk} + \frac{\ell}{2\varphi}(\delta^i_k \partial_j \varphi + \delta^i_j \partial_k \varphi - g_{jk} D^i \varphi) \ ,
\end{aligned} \tag{B.0.2}$$

where D denotes the covariant derivative of g. Equation (B.0.2) can be rewritten as

$$\tilde{D}_X Y = D_X Y + C(X,Y) \ , \tag{B.0.3}$$

with

$$C(X,Y) = \frac{\ell}{2\varphi}\Big(Y(\varphi)X + X(\varphi)Y - g(X,Y)D\varphi\Big) \tag{B.0.4a}$$

$$= \frac{\ell}{2\varphi}\Big(Y(\varphi)X + X(\varphi)Y - \tilde{g}(X,Y)\tilde{D}\varphi\Big) \ . \tag{B.0.4b}$$

B.1 The curvature

Let $\widetilde{\text{Riem}}$ denote the curvature tensor of a connection of the form (B.0.3); from (B.0.4) we obtain

$$\begin{aligned}
\tilde{R}(X,Y)Z &= \Big(\tilde{D}_X \tilde{D}_Y Z - X \leftrightarrow Y\Big) - \tilde{D}_{[X,Y]} Z \\
&= \Big(D_X(D_Y Z + C(Y,Z)) + C(X,(D_Y Z + C(Y,Z)) - X \leftrightarrow Y\Big)
\end{aligned}$$

$$
\begin{aligned}
& -D_{[X,Y]}Z - C(\underbrace{[X,Y]}_{=D_XY-D_YX},Z) \\
= \quad & R(X,Y)Z + \Big((D_XC)(Y,Z) + C(D_XY,Z) + C(Y,D_XZ) \\
& +C(X,D_YZ) + C(X,C(Y,Z)) - X \leftrightarrow Y\Big) - C(D_XY,Z) + C(D_YX,Z) \\
= \quad & R(X,Y)Z + \Big((D_XC)(Y,Z) + C(X,C(Y,Z)) - X \leftrightarrow Y\Big) .
\end{aligned}
$$

In index notation this can be rewritten as

$$\tilde{R}^i{}_{jk\ell} = R^i{}_{jk\ell} + C^i{}_{\ell j;k} - C^i{}_{kj;\ell} + C^i{}_{km}C^m{}_{j\ell} - C^i{}_{\ell m}C^m{}_{jk} \ . \tag{B.1.1}$$

B.1.1 The Weyl conformal connection

There is a natural generalisation of (B.0.3)-(B.0.4) to *Weyl conformal connections*, obtained by the replacement

$$\frac{\ell\partial_a\varphi}{2\varphi} \ \longrightarrow \ f_a \tag{B.1.2}$$

there, where $f_a dx^a$ is an arbitrary one-form, not necessarily exact (compare [71]). In other words, one sets

$$C^i{}_{jk} = \delta^i_j f_k + \delta^i_k f_j - g^{i\ell} f_\ell \, g_{jk} \ . \tag{B.1.3}$$

Since $C^i{}_{jk}$ is symmetric in j and k, the connection $\tilde{D}$ is always torsion-free.

Inserting into (B.1.1) one finds the following formula for the curvature tensor of a Weyl connection

$$\tilde{R}^i{}_{jk\ell} = R^i{}_{jk\ell} + 2\Big(f_{j;[k}\delta^i_{\ell]} + \delta^i_j f_{[\ell;k]} - f^i{}_{;[k}g_{\ell]j} + \delta^i_{[k}f_{\ell]}f_j - g_{j[k}f_{\ell]}f^i - \delta^i_{[k}g_{\ell]j}f_mf^m\Big) \ . \tag{B.1.4}$$

Contracting over i and k one obtains the Ricci tensor of $\tilde{D}$

$$
\begin{aligned}
\tilde{R}_{j\ell} \quad &:= \quad \mathrm{Ric}(\tilde{g})_{ij} \\
&= \quad R_{j\ell} + (1-n)f_{j;\ell} + f_{\ell;j} - f^i{}_{;i}g_{j\ell} + (n-2)(f_jf_\ell - g_{j\ell}f_mf^m) \ .
\end{aligned} \tag{B.1.5}
$$

(Note that $\tilde{R}_{j\ell}$ is not symmetric in general.) We calculate the Ricci scalar of the Weyl connection by taking the trace of $\tilde{R}_{j\ell}$ using the metric g:

$$g^{j\ell}\tilde{R}_{j\ell} \quad = \quad R - (n-1)(2f^i{}_{;i} + (n-2)f_mf^m) \tag{B.1.6}$$

(the reader is warned that this is *not* the curvature scalar of the metric $\tilde{g}$ when f_a is expressed in terms of φ using (B.1.2), see (B.1.14) below).

For $n \neq 2$ it is convenient to introduce the tensor

$$\tilde{L}_{ij} = \frac{1}{n-2}\left(\tilde{R}_{(ij)} - \frac{n-2}{n}\,\tilde{R}_{[ij]} - \frac{1}{2\,(n-1)}\,g_{ij}\,g^{kl}\tilde{R}_{kl}\right) \ . \tag{B.1.7}$$

which is a natural generalisation of the *Schouten tensor* A_{ij} associated to a metric g:

$$A_{ij} = \frac{1}{n-2}\left(R_{ij} - \frac{1}{2(n-1)} g_{ij} g^{k\ell} R_{k\ell}\right) . \tag{B.1.8}$$

From (B.1.5)-(B.1.8) one finds

$$D_i f_j - f_i f_j + \frac{1}{2} g_{ij} f_k f^k = A_{ij} - \tilde{L}_{ij} . \tag{B.1.9}$$

B.1.2 The Weyl tensor

Using (B.1.9) to eliminate the derivatives of f_i from (B.1.4) one obtains

$$\tilde{R}^i{}_{jk\ell} = 2\{\delta^i_{[k}\tilde{L}_{\ell]j} - \delta^i_j \tilde{L}_{[k\ell]} - g_{j[k}\tilde{L}_{\ell]}{}^i\} + C^i{}_{jk\ell} , \tag{B.1.10}$$

where the *Weyl tensor* $C^i{}_{jk\ell}$ is defined as

$$\begin{aligned} C_{ijkl} &:= R_{ijkl} - \frac{1}{n-2}\left(g_{ik}R_{jl} - g_{il}R_{jk} - g_{jk}R_{il} + g_{jl}R_{ik}\right) \\ &\quad + \frac{1}{(n-1)(n-2)} R(g_{ik}g_{jl} - g_{il}g_{jk}) \\ &= R_{ijkl} - A_{ik}g_{jl} + A_{il}g_{jk} + A_{jk}g_{il} - A_{jl}g_{ik} . \end{aligned} \tag{B.1.11}$$

The Weyl tensor has the important property that all its traces vanish, in particular

$$C^i{}_{jik} = 0 .$$

B.1.3 The Ricci tensor and the curvature scalar

We now return to (B.0.4); in this case $\tilde{R}_{ij}$ is the Ricci tensor of the metric $\tilde{g}_{ij}$, hence $\tilde{L}_{ij} = \tilde{A}_{ij}$, the Schouten tensor of $\tilde{g}_{ij}$. Equation (B.1.10) is the statement that $C^i{}_{jk\ell}$ is invariant under conformal changes of the metric:

$$\tilde{C}^i{}_{jk\ell} = C^i{}_{jk\ell} .$$

Next, (B.1.9) can be viewed as a transformation law of the Schouten tensor under conformal changes. Indeed, expressing f_a in terms of φ by inverting (B.1.2), Equation (B.1.9) can be rewritten as

$$\tilde{A}_{ij} = A_{ij} - \frac{\ell}{2\varphi} D_i D_j \varphi + \frac{\ell}{4\varphi^2}\left((2+\ell) D_i\varphi D_j\varphi - \frac{\ell}{2} g_{ij} D_k\varphi D^k\varphi\right) , \tag{B.1.12}$$

which does not have any dimension-dependent coefficients, and which simplifies somewhat when $\ell = -2$. Similarly, (B.1.5) gives

$$\begin{aligned} \tilde{R}_{ij} &= R_{ij} - \frac{(n-2)\ell}{2\varphi} D_i D_j \varphi + \frac{(n-2)\ell(\ell+2)}{4\varphi^2} D_i\varphi D_j\varphi - \frac{\ell}{2\varphi}\Delta_g \varphi g_{ij} \\ &\quad - \frac{(n-2)\ell^2 - 2\ell}{4\varphi^2} D^k\varphi D_k\varphi g_{ij} . \end{aligned} \tag{B.1.13}$$

Taking a $\tilde{g}$–trace one obtains

$$\begin{aligned}\tilde{R} &:= \tilde{g}^{ij}\tilde{R}_{ij} = \varphi^{-\ell} g^{ij}\tilde{R}_{ij} \\ &= \varphi^{-\ell}\left(R - \frac{(n-1)\ell}{\varphi}\Delta_g\varphi - \frac{(n-1)\ell\{(n-2)\ell-4\}}{4\varphi^2}D^i\varphi D_i\varphi\right) .\end{aligned} \tag{B.1.14}$$

For $n \neq 2$ a very convenient choice is

$$(n-2)\ell = 4 , \tag{B.1.15}$$

leading to

$$\tilde{g}_{ij} = \varphi^{\frac{4}{n-2}} g_{ij} , \qquad \tilde{R} = \varphi^{-\frac{4}{n-2}}\left(R - \frac{4(n-1)}{(n-2)\varphi}\Delta_g\varphi\right) . \tag{B.1.16}$$

An immediate useful consequence of (B.1.16) is the following: if $R = 0$ and if φ is g–harmonic (*i.e.*, $\Delta_g\varphi = 0$), then $\tilde{g}$ also has vanishing scalar curvature, and φ is $\tilde{g}$–harmonic.

For $n = 2$ a clever choice is to take $\ell = 1$, and set $\varphi = e^u$, which leads to

$$\tilde{g}_{ij} = e^u g_{ij} , \qquad \tilde{R} = e^{-u}(R - \Delta_g u) . \tag{B.1.17}$$

For the record we note the metric version of (B.1.10),

$$R^i{}_{jk\ell} = 2\{\delta^i_{[k}A_{\ell]j} - g_{j[k}A_{\ell]}{}^i\} + C^i{}_{jk\ell} . \tag{B.1.18}$$

B.2 The wave equation

Under a conformal transformation as in (B.1.16) we have the following transformation law for the Laplacian acting on functions:

$$\begin{aligned}\Delta_{\tilde{g}} f &= \frac{1}{\sqrt{\det \tilde{g}_{ij}}}\partial_k\left(\sqrt{\det \tilde{g}_{ij}}\,\tilde{g}^{k\ell}\partial_\ell f\right) \\ &= \frac{\varphi^{-\frac{2n}{n-2}}}{\sqrt{\det g_{ij}}}\partial_k\Big(\underbrace{\varphi^{\frac{2n}{n-2}-\frac{4}{n-2}}}_{\varphi^2}\sqrt{\det g_{ij}}\,g^{k\ell}\partial_\ell f\Big) \\ &= \varphi^{-\frac{4}{n-2}}(\Delta_g f + 2\varphi^{-1}g^{k\ell}\partial_k\varphi\partial_\ell f) .\end{aligned}$$

This implies

$$\begin{aligned}&\left(\Delta_{\tilde{g}} - \frac{(n-2)}{4(n-1)}\tilde{R}\right) f \\ &= \varphi^{-\frac{4}{n-2}}\left(\Delta_g f + 2\varphi^{-1}g^{k\ell}\partial_k\varphi\partial_\ell f - \frac{(n-2)}{4(n-1)}Rf + \varphi^{-1}f\Delta_g\varphi\right) \\ &= \varphi^{-\frac{4}{n-2}-1}\left(\Delta_g(f\varphi) - \frac{(n-2)}{4(n-1)}Rf\varphi\right) .\end{aligned}$$

Hence the operator

$$\Delta_g - \frac{(n-2)}{4(n-1)}R$$

is conformally-covariant: if $\tilde{g}_{ij} = \varphi^{\frac{4}{n-2}} g_{ij}$, then

$$\left(\Delta_{\tilde{g}} - \frac{(n-2)}{4(n-1)}\tilde{R}\right) f = \varphi^{-\frac{n+2}{n-2}} \left(\Delta_g - \frac{(n-2)}{4(n-1)}R\right) f\varphi \; ; \qquad \text{(B.2.1)}$$

equivalently

$$\left(\Delta_g - \frac{(n-2)}{4(n-1)}R\right) h = \varphi^{\frac{n+2}{n-2}} \left(\Delta_{\tilde{g}} - \frac{(n-2)}{4(n-1)}\tilde{R}\right) \left(\frac{h}{\varphi}\right) \; . \qquad \text{(B.2.2)}$$

B.3 The Cotton tensor

Given any pseudo-Riemannian metric g_{ij}, the Cotton tensor B_{ijk} is defined as

$$B_{ijk} = A_{i[j;k]} \; , \qquad \text{(B.3.1)}$$

where A_{ij} is the Schouten tensor (B.1.8). The tensor B_{ijk} has the following properties

$$\underbrace{B_{ijk} = B_{i[jk]}}_{(a)} \; , \quad \underbrace{B^i{}_{ik}}_{(b)} = 0 \; , \quad \underbrace{B_{[ijk]}}_{(c)} = 0 \; , \qquad \text{(B.3.2)}$$

which, from a purely algebraic point of view, allows a five-dimensional vector space of such tensors at each space point. (The first property in (B.3.2) follows immediately from the definition; similarly the last one is obvious in view of the symmetry of A_{ij} in its indices. The middle-one coincides with the contracted Bianchi identity, $R_i{}^j{}_{;j} = \frac{1}{2}R_{;j}$.)

The Cotton tensor further satisfies the differential identity

$$B_{i[jk;l]} = 0 \; . \qquad \text{(B.3.3)}$$

One can think of the Cotton tensor as the three-dimensional counterpart of the Weyl tensor. Indeed, the Weyl tensor vanishes identically in dimension three so it is not of much interest there. On the other hand, B transforms homogeneously under conformal transformation when $n = 3$. Indeed,

In dimension three, an object equivalent to the Cotton tensor is the tensor

$$H_{ij} = \frac{1}{2}\epsilon^{kl}{}_i B_{jkl} \; . \qquad \text{(B.3.4)}$$

The tensor H_{ij} is symmetric, tracefree and divergence-free. Indeed, the vanishing of its trace is precisely (B.3.2)(c). The vanishing of the divergence is (B.3.3). To see the symmetry, we calculate as follows

$$H_{12} \underbrace{=}_{\text{def.}} B_{223} = \underbrace{B_{113} + B_{223} + B_{333}}_{=0 \text{ by (B.3.2)(b)}} - B_{113} - \underbrace{B_{333}}_{=0 \text{ by (B.3.2)(a)}} = - \underbrace{B_{113}}_{=-B_{131} \text{ by (B.3.2)(a)}}$$

$$\underbrace{=}_{\text{def.}} H_{21} \; .$$

Finally, one readily verifies the inversion formula

$$B_{ijk} = \epsilon_{jk}{}^{\ell} H_{i\ell} \; . \qquad \text{(B.3.5)}$$

B.4 The Bach tensor

The Bach tensor is defined by the formula

$$B_{ab} = D^c D^d C_{abcd} + \frac{1}{2} R^{cd} C_{acbd} \tag{B.4.1}$$

Its interest arises from the fact that it is conformally covariant in four dimensions,

$$g_{ij} \to \omega^2 g_{ij} \quad \Longrightarrow \quad B_{ij} \to \omega^{-2} B_{ij} \ .$$

Whatever the dimension, B_{ij} vanishes if g is Einstein. This follows from the fact that $D^d C_{abcd}$ vanishes for Einstein metrics by the Bianchi identity, while the second term in (B.4.1) becomes a trace in the second and third index, which is zero for the Weyl tensor.

B.5 The Graham-Hirachi theorem, and the Fefferman-Graham obstruction tensor

B.5.1 The Fefferman-Graham tensor

Let, as elsewhere, $n+1$ denote space-time dimension, with n odd. The Fefferman-Graham tensor $\mathcal{H}$ is a conformally covariant tensor, built out of the metric g and its derivatives up to order $n+1$, of the form

$$\mathcal{H} = (\nabla^*\nabla)^{\frac{n+1}{2}-2}[\nabla^*\nabla(A) + \nabla^2(trA)] + \mathcal{F}^n \ , \tag{B.5.1}$$

where A is the Schouten tensor (B.1.8), and where $\mathcal{F}^n$ is a tensor built out of lower order derivatives of the metric (see, e.g., [79], where the notation $\mathcal{O}$ is used in place of $\mathcal{H}$). It turns out that $\mathcal{F}^n$ involves only derivatives of the metric up to order $n-1$: this is an easy consequence of Equation (2.4) in [79], using the fact that odd-power coefficients of the expansion of the metric g_x in [79, Equation (2.3)] vanish. (For $n = 3, 5$ this can also be verified by inspection of the explicit formulae for $\mathcal{F}^3$ and $\mathcal{F}^5$ given in [79].)

The system of equations

$$\mathcal{H} = 0 \tag{B.5.2}$$

will be called the Anderson-Fefferman-Graham (AFG) equations. It has the following properties [79]:

1. The system (B.5.2) is conformally invariant: if g is a solution, so is $\varphi^2 g$, for any positive function φ.

2. If g is conformal to an Einstein metric, then (B.5.2) holds.

3. $\mathcal{H}$ is trace-free.

4. $\mathcal{H}$ is divergence-free.

The tensor $\mathcal{H}$ was originally discovered by Fefferman and Graham [63] as an obstruction to the existence of a formal power series expansion for conformally compactifiable Einstein metrics, with conformal boundary equipped with the conformal equivalence class $[g]$ of g. This geometric interpretation is irrelevant from our point of view, as here we are interested in (B.5.2) as an equation on its own.

B.5.2 The Graham-Hirachi theorem

It is of interest to classify all conformally-covariant tensors which are polynomial in the metric, its inverse, and in the derivatives of the metric. Such tensors will be called natural. Now, one may construct further covariants from known ones by taking tensor products and contracting. A tensor will be called irreducible if it cannot be constructed in that fashion in a non-trivial way.

The following theorem of Hirachi-Graham shows that up to quadratic and higher terms in curvature, the Weyl tensor, or the Cotton tensor in dimension 3, and the obstruction tensor are the only irreducible conformally invariant tensors:

THEOREM B.5.1 (Graham-Hirachi [79]) *A conformally covariant irreducible natural tensor of n-dimensional oriented Riemannian manifolds is equivalent modulo a conformally covariant natural tensor of degree at least 2 in curvature with a multiple of one of the following:*

1. $n = 3$*: the Cotton tensor* $C_{ijk} = A_{ij;k} - A_{ik;j}$
2. $n = 4$*: the self-dual or anti-self dual Weyl tensor* $C^{\pm}_{ijkl}$ *or the Bach tensor* $B_{ij} = \mathcal{O}_{ij}$
3. $n \geq 5$ *odd: the Weyl tensor* C_{ijkl}
4. $n \geq 6$ *even: the Weyl tensor* C_{ijkl} *or the obstruction tensor* $\mathcal{O}_{ij}$

B.6 Frame coefficients, Dirac operators

In order to calculate the transformation law of the connection coefficients, we will consider a conformal rescaling of the form $\bar g_{ij} = e^{2u} g_{ij}$. Let $\bar\theta^i$ be an orthonormal coframe for $\bar g$, then

$$\theta^i := e^{-u}\bar\theta^i$$

is an orthonormal coframe for $\bar g$. We claim that:

$$\bar\omega_{ij}(e_k) = \omega_{ij}(e_k) - e_i(u)g_{jk} + e_j(u)g_{ik} \ , \tag{B.6.1}$$

equivalently

$$\bar\omega_{ij} = \bar\omega_{ij}(e_k)\theta^k = \omega_{ij} - e_i(u)\theta_j + e_j(u)\theta_i \ . \tag{B.6.2}$$

To verify this equation, notice that $\bar\omega_{ij}$ as given by this equation is anti-symmetric in i and j; further, it is straightforward to check that

$$d\bar\theta^i + \bar\omega^i{}_j \wedge \bar\theta^j = 0 \ ,$$

and (B.6.1) follows from uniqueness of $\bar\omega_{ij}$.

Let e_i be an orthonormal frame for g. Recall that the Dirac operator Dirac is defined by the formula

$$\mathrm{Dirac}\psi := \gamma^k \nabla_{e^k}\psi = \gamma^k(e_k(\psi) - \frac{1}{4}\omega_{ijk}\gamma^i\gamma^j)\psi \;.$$

The corresponding Dirac operator $\overline{\mathrm{Dirac}}$ associated to the metric $\bar g$ reads

$$\overline{\mathrm{Dirac}}\psi := \gamma^k \bar\nabla_{\bar e^k}\psi = \gamma^k(\bar e_k(\psi) - \frac{1}{4}\bar\omega_{ijk}\gamma^i\gamma^j)\psi \;.$$

Using (B.6.1) one finds

$$\overline{\mathrm{Dirac}}\psi = e^{-\frac{(n+1)u}{2}}\mathrm{Dirac}(e^{\frac{(n-1)u}{2}}\psi) \;. \tag{B.6.3}$$

B.7 Non-characteristic hypersurfaces

Let $\mathscr{S}$ be a non-characteristic hypersurface in $\mathscr{M}$, under (B.0.1) the unit normal to $\mathscr{S}$ transforms as

$$\tilde n^i = \varphi^{-\ell/2} n^i \quad \Longleftrightarrow \quad \tilde n_i = \varphi^{\ell/2} n_i \;. \tag{B.7.1}$$

The projection tensor P defined in (1.3.4) is invariant under (B.0.1),

$$\tilde P = P \;.$$

From the definition (1.3.5) of the Weingarten map we obtain, for $X \in T\mathscr{S}$,

$$\begin{aligned}
\tilde B(X) &= \tilde P(\tilde D_X \tilde n) = P\Big(D_X(\varphi^{-\ell/2}n) + C(X,\tilde n)\Big)\\
&= P\Big(X(\varphi^{-\ell/2})n + \varphi^{-\ell/2}D_X n + \frac{\ell}{2\varphi}\Big(\tilde n(\varphi)X + X(\varphi)\tilde n - \underbrace{\tilde g(X,\tilde n)}_{0}\tilde D\varphi\Big)\\
&= P\Big(\varphi^{-\ell/2}D_X n + \frac{\ell}{2\varphi}\tilde n(\varphi)X\Big)\\
&= \varphi^{-\ell/2}B(X) + \frac{\ell}{2\varphi}\tilde n(\varphi)X \;.
\end{aligned}$$

The definition (1.3.6) of the extrinsic curvature tensor (second fundamental form) K leads to, for $X, Y \in T\mathscr{S}$,

$$\tilde K(X,Y) = \tilde g\Big(\tilde B(X), Y\Big) = \tilde g\Big(\varphi^{-\ell/2}B(X) + \frac{\ell}{2\varphi}\tilde n(\varphi)X, Y\Big) \;,$$

which can be rewritten in the following three equivalent forms

$$\tilde K(X,Y) = \varphi^{\ell/2}K(X,Y) + \frac{\ell\tilde n(\varphi)}{2\varphi}\tilde g(X,Y) \tag{B.7.2a}$$

$$= \varphi^{\ell/2}K(X,Y) + \frac{\ell\varphi^{(-\ell-2)/2}n(\varphi)}{2}\tilde g(X,Y) \tag{B.7.2b}$$

$$= \varphi^{\ell/2}K(X,Y) + \frac{\ell\varphi^{(\ell-2)/2}n(\varphi)}{2}g(X,Y) \;. \tag{B.7.2c}$$

Bibliography

[1] R.A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[2] S. Agmon, A. Douglis, and L. Nirenberg, *Estimates near the boundary for the solutions of elliptic equations satisfying general boundary values*, Commun. Pure Appl. Math.**12** (1959), 623–727.

[3] P. Allen, A. Clausen, and J. Isenberg, *Near-constant mean curvature solutions of the Einstein constraint equations with non-negative Yamabe metrics*, Class. Quantum Grav. **25** (2008), 075009, 15, arxiv:0710.0725v1[gr-qc]. MR MR2404418

[4] L. Andersson and P.T. Chruściel, *On "hyperboloidal" Cauchy data for vacuum Einstein equations and obstructions to smoothness of null infinity*, Phys. Rev. Lett. **70** (1993), 2829–2832. MR MR1215407 (94a:83020)

[5] ______, *On asymptotic behavior of solutions of the constraint equations in general relativity with "hyperboloidal boundary conditions"*, Dissert. Math. **355** (1996), 1–100. MR MR1405962 (97e:58217)

[6] L. Andersson, P.T. Chruściel, and H. Friedrich, *On the regularity of solutions to the Yamabe equation and the existence of smooth hyperboloidal initial data for Einsteins field equations*, Commun. Math. Phys. **149** (1992), 587–612. MR MR1186044 (93i:53040)

[7] L. Andersson and V. Moncrief, *Elliptic-hyperbolic systems and the Einstein equations*, Ann. Henri Poincaré **4** (2003), 1–34, arXiv:gr-qc/0110111.

[8] T. Aubin, *Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire*, Jour. Math. Pures Appl. (9) **55** (1976), 269–296. MR 0431287 (55 #4288)

[9] ______, *Espaces de Sobolev sur les variétés Riemanniennes*, Bull. Sci. Math., II. Ser. **100** (1976), 149–173.

[10] ______, *Nonlinear analysis on manifolds. Monge–Ampère equations*, Springer, New York, Heidelberg, Berlin, 1982.

[11] A. Bahri, *Critical points at infinity in some variational problems*, Pitman Research Notes in Mathematics Series, vol. 182, Longman Scientific & Technical, Harlow, 1989. MR 1019828 (91h:58022)

[12] R. Bartnik, *The existence of maximal hypersurfaces in asymptotically flat space-times*, Commun. Math. Phys. **94** (1984), 155–175.

[13] ______, *The mass of an asymptotically flat manifold*, Commun. Pure Appl. Math. **39** (1986), 661–693.

[14] ______, *Regularity of variational maximal surfaces*, Acta Math. **161** (1988), 145–181.

[15] ______, *Remarks on cosmological spacetimes and constant mean curvature surfaces*, Commun. Math. Phys. **117** (1988), 615–624.

[16] ______, *Some open problems in mathematical relativity*, Conference on Mathematical Relativity (Canberra, 1988), Proc. Centre Math. Anal. Austral. Nat. Univ., vol. 19, Austral. Nat. Univ., Canberra, 1989, pp. 244–268. MR MR1020805 (90g:83001)

[17] ______, *Quasi-spherical metrics and prescribed scalar curvature*, Jour. Diff. Geom. **37** (1993), 31–71. MR 93i:53041

[18] ______, *Phase space for the Einstein equations*, Commun. Anal. Geom. **13** (2005), 845–885, arXiv:gr-qc/0402070. MR MR2216143 (2007d:83012)

[19] R. Bartnik and G. Fodor, *On the restricted validity of the thin sandwich conjecture*, Phys. Rev. D **48** (1993), 3596–3599.

[20] R. Bartnik and J. Isenberg, *The constraint equations*, The Einstein equations and the large scale behavior of gravitational fields, Birkhäuser, Basel, 2004, pp. 1–38. MR MR2098912 (2005j:83007)

[21] T.W. Baumgarte and S.L. Shapiro, *Numerical integration of Einstein's field equations*, Phys. Rev. D (3) **59** (1999), 024007, 7, arXiv:gr-qc/9810065. MR MR1692065 (2000b:83001)

[22] R. Beig, *TT-tensors and conformally flat structures on* 3*-manifolds*, Mathematics of gravitation, Part I (Warsaw, 1996), Banach Center Publ., vol. 41, Polish Acad. Sci., Warsaw, 1997, pp. 109–118. MR MR1466511 (98k:53040)

[23] R. Beig and P.T. Chruściel, *Killing Initial Data*, Class. Quantum. Grav. **14** (1997), A83–A92, A special issue in honour of Andrzej Trautman on the occasion of his 64th Birthday, J.Tafel, editor. MR MR1691888 (2000c:83011)

[24] R. Beig, P.T. Chruściel, and R. Schoen, *KIDs are non-generic*, Ann. H. Poincaré **6** (2005), 155–194, arXiv:gr-qc/0403042. MR MR2121280 (2005m:83013)

[25] A.L. Besse, *Einstein manifolds*, Ergebnisse d. Math. 3. folge, vol. 10, Springer, Berlin, 1987.

[26] L. Bessières, G. Besson, and S. Maillot, *Ricci flow on open 3-manifolds and positive scalar curvature*, (2010), arXiv:1001.1458 [math.dg].

[27] N.T. Bishop, R. Gómez, L. Lehner, M. Maharaj, and J. Winicour, *Characteristic initial data for a star orbiting a black hole*, Phys. Rev. D **72** (2005), 024002, 16. MR MR2171948 (2006d:83051)

[28] J.-P. Bourguignon, D.G. Ebin, and J.E. Marsden, *Sur le noyau des opérateurs pseudo-différentiels à symbole surjectif et non injectif*, C. R. Acad. Sci. Paris Sér. A **282** (1976), 867–870. MR 53 #6643

[29] R. Bousso, *Adventures in de Sitter space*, The future of the theoretical physics and cosmology (Cambridge, 2002), Cambridge Univ. Press, Cambridge, 2003, pp. 539–569. MR 2033285

[30] D. Brill and F. Flaherty, *Maximizing properties of extremal surfaces in general relativity*, Ann. Inst. H. Poincaré Sect. A (N.S.) **28** (1978), 335–347. MR MR0479299 (57 #18741)

[31] G. Caciotta and F. Nicolò, *Global characteristic problem for Einstein vacuum equations with small initial data. I. The initial data constraints*, Jour. Hyp. Differ. Equ. **2** (2005), no. 1, 201–277, arXiv:gr-qc/0409028. MR MR2134959 (2006i:58042)

[32] ______, *On a class of global characteristic problems for the Einstein vacuum equations with small initial data*, Jour. Math. Phys. **51** (2010), 102503 (21 pp.), arXiv: gr-qc/0608038.

[33] M. Cantor, *The existence of non-trivial asymptotically flat initial data for vacuum spacetimes*, Commun. Math. Phys. **57** (1977), 83–96. MR MR0462440 (57 #2414)

[34] Y. Choquet-Bruhat, *Einstein constraints on n dimensional compact manifolds*, Class. Quantum Grav. **21** (2004), S127–S152, arXiv:gr-qc/0311029.

[35] ______, *General relativity and the Einstein equations*, Oxford Mathematical Monographs, Oxford University Press, Oxford, UK, 2009. MR MR2473363 (2010f:83001)

[36] Y. Choquet-Bruhat and D. Christodoulou, *Elliptic systems in $H_{s,\delta}$ spaces on manifolds which are Euclidean at infinity*, Acta Math. **146** (1981), 129–150. MR MR594629 (82c:58060)

[37] Y. Choquet-Bruhat, P.T. Chruściel, and J.M. Martín-García, *The Cauchy problem on a characteristic cone for the Einstein equations in arbitrary dimensions*, Ann. H. Poincaré (2010), in press, arXiv:1006.4467 [gr-qc].

[38] Y. Choquet-Bruhat, J. Isenberg, and V. Moncrief, *Solutions of constraints for Einstein equations*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), 349–355. MR MR1179734 (93h:58151)

[39] Y. Choquet-Bruhat, J. Isenberg, and D. Pollack, *The Einstein-scalar field constraints on asymptotically Euclidean manifolds*, Chinese Ann. Math. Ser. B **27** (2006), 31–52, arXiv:gr-qc/0506101. MR MR2209950 (2007h:58049)

[40] ———, *Applications of theorems of Jean Leray to the Einstein-scalar field equations*, Jour. Fixed Point Theory Appl. **1** (2007), 31–46, arXiv:gr-qc/0611009. MR MR2282342

[41] ———, *The constraint equations for the Einstein-scalar field system on compact manifolds*, Class. Quantum Grav. **24** (2007), 809–828, arXiv:gr-qc/0610045. MR MR2297268

[42] Y. Choquet-Bruhat, J. Isenberg, and J.W. York, Jr., *Einstein constraints on asymptotically Euclidean manifolds*, Phys. Rev. D **61** (2000), 084034 (20 pp.), arXiv:gr-qc/9906095.

[43] Y. Choquet-Bruhat and J. York, *The Cauchy problem*, General Relativity (A. Held, ed.), Plenum Press, New York, 1980, pp. 99–172.

[44] D. Christodoulou, *The Formation of Black Holes in General Relativity*, (2008), x+589, arXiv:0805.3880 [gr-qc]. MR MR2488976 (2009k:83010)

[45] D. Christodoulou and N.Ó Murchadha, *The boost problem in general relativity*, Commun. Math. Phys. **80** (1981), 271–300.

[46] P.T. Chruściel, J. Corvino, and J. Isenberg, *Construction of N-body time-symmetric initial data sets in general relativity*, (2009), arXiv:0909.1101 [gr-qc].

[47] ———, *Construction of N-body initial data sets in general relativity*, (2010), arXiv:0909.1101 [gr-qc].

[48] P.T. Chruściel and E. Delay, *Existence of non-trivial asymptotically simple vacuum space-times*, Class. Quantum Grav. **19** (2002), L71–L79, arXiv:gr-qc/0203053, erratum-ibid, 3389. MR MR1902228 (2003e:83024a)

[49] ———, *On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications*, Mém. Soc. Math. de France. **94** (2003), vi+103, arXiv:gr-qc/0301073v2. MR MR2031583 (2005f:83008)

[50] ———, *Manifold structures for sets of solutions of the general relativistic constraint equations*, Jour. Geom Phys. (2004), 442–472, arXiv:gr-qc/0309001v2. MR MR2085346 (2005i:83008)

[51] ———, *Gluing constructions for asymptotically hyperbolic manifolds with constant scalar curvature*, (2007), arXiv:0710.xxx [gr-qc].

[52] P.T. Chruściel, J. Isenberg, and D. Pollack, *Gluing initial data sets for general relativity*, Phys. Rev. Lett. **93** (2004), 081101, arXiv:gr-qc/0409047.

[53] ———, *Initial data engineering*, Commun. Math. Phys. **257** (2005), 29–42, arXiv:gr-qc/0403066. MR MR2163567 (2007d:83013)

[54] P.T. Chruściel, F. Pacard, and D. Pollack, *Singular Yamabe metrics and initial data with exactly Kottler-Schwarzschild-de Sitter ends II. Generic metrics*, Math. Res. Lett. **16** (2008), 157–164, arXiv:0803.1817 [gr-qc]. MR 2480569 (2009k:53079)

[55] P.T. Chruściel and D. Pollack, *Singular Yamabe metrics and initial data with* exactly *Kottler–Schwarzschild–de Sitter ends*, Ann. Henri Poincaré (2008), 639–654, arXiv:0710.3365 [gr-qc]. MR 2413198 (2009g:53051)

[56] J. Corvino, *Scalar curvature deformation and a gluing construction for the Einstein constraint equations*, Commun. Math. Phys. **214** (2000), 137–189. MR MR1794269 (2002b:53050)

[57] ———, *On the existence and stability of the Penrose compactification*, Annales H. Poincaré **8** (2007), 597–620. MR MR2329363

[58] J. Corvino and R.M. Schoen, *On the asymptotics for the vacuum Einstein constraint equations*, Jour. Diff. Geom. **73** (2006), 185–217, arXiv:gr-qc/0301071. MR MR2225517 (2007e:58044)

[59] S. Dain, *Trapped surfaces as boundaries for the constraint equations*, Class. Quantum Grav. **21** (2004), 555–574, Corrigendum ib. p. 769, arXiv:gr-qc/0308009.

[60] S. Dain and H. Friedrich, *Asymptotically flat initial data with prescribed regularity at infinity*, Commun. Math. Phys. **222** (2001), no. 3, 569–609. MR MR1888089 (2003f:58057)

[61] C. Delaunay, *Sur la surface de révolution dont la courbure moyenne est constante*, Jour. de Math. **6** (1841), 309–320.

[62] M. Dossa, *Problèmes de Cauchy sur un conoïde caractéristique pour les équations d'Einstein (conformes) du vide et pour les équations de Yang-Mills-Higgs*, Ann. Henri Poincaré **4** (2003), 385–411. MR MR1985778 (2004h:58041)

[63] C. Fefferman and C.R. Graham, *Conformal invariants*, Élie Cartan et les mathématiques d'aujourd'hui, The mathematical heritage of Élie Cartan, Sémin. Lyon 1984, Astérisque, No.Hors Sér. 1985, 95-116.

[64] R. Fintushel, *Circle actions on simply connected 4-manifolds.*, Trans. Am. Math. Soc. **230** (1977), 147–171.

[65] ———, *Classification of circle actions on 4-manifolds.*, Trans. Am. Math. Soc. **242** (1978), 377–390.

[66] A. Fischer, *The theory of superspace*, Relativity: Proceedings of the relativity conference in the Midwest (M. Carmeli, S.I. Fickler, and L. Witten, eds.), Plenum, New York, 1970.

[67] A. Fischer, J.E. Marsden, and V. Moncrief, *The structure of the space of solutions of Einstein's equations. I. One Killing field*, Ann. Inst. H. Poincaré Sect. A (N.S.) **33** (1980), no. 2, 147–194. MR 82d:83007

[68] Y. Fourès-Bruhat, *Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires*, Acta Math. **88** (1952), 141–225.

[69] R.H. Fowler, *The form near infinity of real continuous solutions of a certain differential equation of the second order.*, Quart. Jour. Pure Appl. Math. **45** (1914), 289–349.

[70] ———, *Further studies of Emden's and similar differential equations.*, Quart. Jour. Math., Oxf. Ser. **2** (1931), 259–288.

[71] H. Friedrich, *Conformal Einstein evolution*, Proceedings of the Tübingen Workshop on the Conformal Structure of Space-times, H. Friedrich and J. Frauendiener, Eds., Springer Lecture Notes in Physics **604**, 1–50 (2002), arXiv:gr-qc/0209018.

[72] ———, *Hyperbolic reductions for Einstein's equations*, Class. Quantum Grav. **13** (1996), 1451–1469.

[73] ———, *Yamabe numbers and the Brill-Cantor criterion*, (2004), preprint.

[74] H. Friedrich and G. Nagy, *The initial boundary value problem for Einstein's vacuum field equation*, Commun. Math. Phys. **201** (1998), 619–655.

[75] H. Friedrich and A. Rendall, *The Cauchy problem for the Einstein equations*, Einstein's field equations and their physical implications, Lecture Notes in Phys., vol. 540, Springer, Berlin, 2000, pp. 127–223. MR MR1765130 (2001m:83009)

[76] H. Friedrich and A.D Rendall, *The Cauchy problem for the Einstein equations*, Einstein's field equations and their physical implications, Lecture Notes in Phys., vol. 540, Springer, Berlin, 2000, arXiv:gr-qc/0002074, pp. 127–223. MR MR1765130 (2001m:83009)

[77] C. Gerhardt, *H-surfaces in Lorentzian manifolds*, Commun. Math. Phys. **89** (1983), 523–553. MR MR713684 (85h:53049)

[78] R. Glassey and M. Tsutsumi, *On uniqueness of weak solutions to semilinear wave equations*, Commun. Partial Differential Equations **7** (1982), no. 2, 153–195.

[79] C.R. Graham and K. Hirachi, *The ambient obstruction tensor and Q-curvature*, AdS/CFT correspondence: Einstein metrics and their conformal boundaries, IRMA Lect. Math. Theor. Phys., vol. 8, Eur. Math. Soc., Zürich, 2005, pp. 59–71.

[80] M. Gromov and H.B. Lawson, Jr., *Spin and scalar curvature in the presence of a fundamental group. I*, Ann. of Math. (2) **111** (1980), 209–230. MR MR569070 (81g:53022)

[81] ———, *Positive scalar curvature and the Dirac operator on complete Riemannian manifolds*, Inst. Hautes Études Sci. Publ. Math. (1983), 83–196. MR MR720933 (85g:58082)

[82] E. Hebey, *Existence, stability and instability for Einstein–scalar field Lichnerowicz equations*, 2009, `http://www.u-cergy.fr/rech/pages/hebey/IASBeamerFullPages.pdf`.

[83] E. Hebey, F. Pacard, and D. Pollack, *A variational analysis of Einstein-scalar field Lichnerowicz equations on compact Riemannian manifolds*, Commun. Math. Phys. **278** (2008), no. 1, 117–132, arXiv:gr-qc/0702031. MR MR2367200

[84] M. Holst, L. Lindblom, R. Owen, H.P. Pfeiffer, M.A. Scheel, and L.E. Kidder, *Optimal constraint projection for hyperbolic evolution systems*, Phys. Rev. D (3) **70** (2004), 084017, 17. MR MR2117121 (2005j:83008)

[85] M. Holst, G. Nagy, and G. Tsogtgerel, *Rough solutions of the Einstein constraint equations on closed manifolds without near-CMC conditions*, Commun. Math. Phys., arXiV:0712.0798 [gr-qc]. MR MR2500992

[86] ———, *Far-from-constant mean curvature solutions of Einstein's constraint equations with positive Yamabe metrics*, Phys. Rev. Lett. **100** (2008), 161101, 4. MR MR2403263

[87] J. Isenberg, *Constant mean curvature solutions of the Einstein constraint equations on closed manifolds*, Class. Quantum Grav. **12** (1995), 2249–2274. MR MR1353772 (97a:83013)

[88] ———, *The constraint problem for Einstein's equations*, Living Reviews in Relativity (1998), URL `http://www.livingreviews.org`.

[89] J. Isenberg, D. Maxwell, and D. Pollack, *Gluing of non-vacuum solutions of the Einstein constraint equations*, Adv. Theor. Math. Phys. **9** (2005), 129–172. MR MR2193370 (2006j:83004)

[90] J. Isenberg, R. Mazzeo, and D. Pollack, *Gluing and wormholes for the Einstein constraint equations*, Commun. Math. Phys. **231** (2002), 529–568, arXiv:gr-qc/0109045.

[91] ———, *Gluing and wormholes for the Einstein constraint equations*, Commun. Math. Phys. **231** (2002), 529–568, arXiv:gr-qc/0109045. MR MR1946448 (2004a:83006)

[92] ———, *On the topology of vacuum spacetimes*, Ann. Henri Poincaré **4** (2003), 369–383. MR MR1985777 (2004h:53053)

[93] J. Isenberg and V. Moncrief, *Some results on non–constant mean curvature solutions of the Einstein constraint equations*, Physics on Manifolds (M. Flato, R. Kerner, and A. Lichnerowicz, eds.), Kluwer Academic Publishers, Dordrecht, 1994, Y. Choquet–Bruhat Festschrift, pp. 295–302.

[94] ______, *A set of nonconstant mean curvature solutions of the Einstein constraint equations on closed manifolds*, Class. Quantum Gravity **13** (1996), 1819–1847. MR MR1400943 (97h:83010)

[95] J. Isenberg and N. Ó Murchadha, *Non-CMC conformal data sets which do not produce solutions of the Einstein constraint equations*, Class. Quantum Grav. **21** (2004), S233–S241, A space-time safari: Essays in honour of Vincent Moncrief; arXiv:gr-qc/0311057. MR MR2053007 (2005c:83003)

[96] J. Isenberg and J. Park, *Asymptotically hyperbolic non-constant mean curvature solutions of the Einstein constraint equations*, Classical Quantum Grav. **14** (1997), A189–A201. MR 2000e:83009

[97] D. Joyce, *Constant scalar curvature metrics on connected sums*, Int. Jour. Math. Math. Sci. (2003), 405–450, arXiv:math.DG/0108022. MR 1 961 016

[98] S. Klainerman and I. Rodnianski, *Ricci defects of microlocalized Einstein metrics*, Jour. Hyperbolic Differ. Equ. **1** (2004), 85–113. MR MR2052472 (2005f:58048)

[99] ______, *The causal structure of microlocalized rough Einstein metrics*, Ann. of Math. (2) **161** (2005), 1195–1243. MR MR2180401 (2007d:58052)

[100] ______, *Rough solutions of the Einstein-vacuum equations*, Ann. of Math. (2) **161** (2005), 1143–1193. MR MR2180400 (2007d:58051)

[101] V.I. Kondrashov, *On certain estimates of families of functions satisfying integral inequalities*, Dokl. Akad. Nauk SSSR **48** (1945), 565–566 (Russian).

[102] N. Korevaar, R. Mazzeo, F. Pacard, and R. Schoen, *Refined asymptotics for constant scalar curvature metrics with isolated singularities*, Invent. Math. **135** (1999), 233–272. MR MR1666838 (2001a:35055)

[103] F. Kottler, *Über die physikalischen Grundlagen der Einsteinschen Gravitationstheorie*, Annalen der Physik **56** (1918), 401–462.

[104] H.O. Kreiss, O. Reula, O. Sarbach, and J. Winicour, *Well-posed initial-boundary value problem for the harmonic Einstein equations using energy estimates*, Class. Quant. Grav. **24** (2007), 5973–5984, arXiv:0707.4188 [gr-qc].

[105] W. Kühnel and H.-B. Rademacher, *Essential conformal fields in pseudo–Riemannian geometry*, Jour. Math. Sci. Univ. Tokyo **4** (1997), 649–662.

[106] J.M. Lee and T.H. Parker, *The Yamabe problem*, Bull. Amer. Math. Soc. (N.S.) **17** (1987), 37–91. MR MR888880 (88f:53001)

[107] J. Lelong-Ferrand, *Geometrical interpretations of scalar curvature and regularity of conformal homeomorphisms* , Differ. Geom. Relativ., Vol. Honour A. Lichnerowicz 60th Birthday, 1976, pp. 91–105.

[108] A. Lichnerowicz, *L'intégration des équations de la gravitation relativiste et le problème des n corps*, Jour. Math. Pures Appli. **23** (1944), 37–63.

[109] R.B. Lockhart, *Fredholm properties of a class of elliptic operators on non-compact manifolds*, Duke Math. J. **48** (1981), 289–312. MR MR610188 (82j:35050)

[110] R.B. Lockhart and R.C. McOwen, *On elliptic systems in $\mathbf{R}^n$*, Acta Math. **150** (1983), 125–135. MR MR697610 (84d:35048)

[111] ———, *Correction to: "On elliptic systems in $\mathbf{R}^n$" [Acta Math.* **150** *(1983), no. 1-2, 125–135]*, Acta Math. **153** (1984), 303–304. MR MR766267 (86a:35049)

[112] ———, *Elliptic differential operators on noncompact manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **12** (1985), 409–447. MR MR837256 (87k:58266)

[113] F.C. Marques, *Isolated singularities of solutions of the Yamabe equation*, Calc. of Var. (2007), doi:10.1007/s00526–007–0144–3.

[114] J.M. Martín-García, *xAct: Efficient Tensor Computer Algebra*, `http://metric.iem.csic.es/Martin-Garcia/xAct`.

[115] D. Maxwell, *Rough solutions of the Einstein constraint equations on compact manifolds*, Jour. Hyperbolic Diff. Equ. **2** (2005), 521–546, arXiv:gr-qc/0506085. MR MR2151120 (2006d:58027)

[116] ———, *Solutions of the Einstein constraint equations with apparent horizon boundaries*, Commun. Math. Phys. **253** (2005), 561–583, arXiv:gr-qc/0307117. MR MR2116728 (2006c:83008)

[117] ———, *Rough solutions of the Einstein constraint equations*, J. Reine Angew. Math. **590** (2006), 1–29, arXiv:gr-qc/0405088. MR MR2208126 (2006j:58044)

[118] ———, *A class of solutions of the vacuum Einstein constraint equations with freely specified mean curvature*, Math. Res. Lett. **16** (2009), 627–645, arXiv:0804.0874v1 [gr-qc]. MR MR2525029

[119] R. Mazzeo and F. Pacard, *Constant scalar curvature metrics with isolated singularities*, Duke Math. J. **99** (1999), 353–418. MR MR1712628 (2000g:53035)

[120] R. Mazzeo, D. Pollack, and K. Uhlenbeck, *Connected sum constructions for constant scalar curvature metrics*, Topol. Methods Nonlinear Anal. **6** (1995), 207–233. MR MR1399537 (97e:53076)

[121] ______, *Moduli spaces of singular Yamabe metrics*, Jour. Amer. Math. Soc. **9** (1996), 303–344. MR MR1356375 (96f:53055)

[122] R.C. McOwen, *The behavior of the Laplacian on weighted Sobolev spaces*, Commun. Pure Appl. Math. **32** (1979), 783–795. MR MR539158 (81m:47069)

[123] ______, *Fredholm theory of partial differential equations on complete Riemannian manifolds*, Pacific J. Math. **87** (1980), 169–185. MR MR590874 (82g:58082)

[124] ______, *On elliptic operators in* $\mathbf{R}^n$, Commun. Partial Differential Equations **5** (1980), 913–933. MR MR584101 (81k:35058)

[125] C.W. Misner, K. Thorne, and J.A. Wheeler, *Gravitation*, Freeman, San Fransisco, 1973.

[126] V. Moncrief, *Spacetime symmetries and linearization stability of the Einstein equations I*, Jour. Math. Phys. **16** (1975), no. 3, 493–498.

[127] C.B. Morrey, *Multiple integrals in the calculus of variation*, Die Grundlehren der mathematischen Wissenschaften, Band 130, Springer Verlag, Berlin, Heidelberg, New York, 1966. MR MR0202511 (34 #2380)

[128] H. Nariai, *On a new cosmological solution of Einstein's field equations of gravitation*, Sci. Rep. Tohoku Univ. Ser. I. **35** (1951), 62–67. MR 14,1133f

[129] L. Nirenberg and H.F. Walker, *The null spaces of elliptic partial differential operators in* $\mathbf{R}^n$, J. Math. Anal. Appl. **42** (1973), 271–301, Collection of articles dedicated to Salomon Bochner. MR MR0320821 (47 #9354)

[130] M. Obata, *The conjectures of conformal transformations of Riemannian manifolds.*, Bull. Amer. Math. Soc. **77** (1971), 265–270. MR MR0270397 (42 #5286)

[131] F. Pacard and T. Rivière, *Linear and nonlinear aspects of vortices*, Progress in Nonlinear Differential Equations and their Applications, 39, Birkhäuser Boston Inc., Boston, MA, 2000, The Ginzburg-Landau model. MR MR1763040 (2001k:35066)

[132] V. Paschalidis, *Mixed Hyperbolic-Second-Order Parabolic Formulations of General Relativity*, (2007), arXiv:0704.2861 [gr-qc].

[133] R. Penrose, *Asymptotic properties of fields and space–times*, Phys. Rev. Lett. **10** (1963), 66–68.

[134] G. Perelman, *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*, (2003), arXiv:math/0307245.

[135] ———, *Ricci flow with surgery on three-manifolds*, (2006), arXiv:math/0303109.

[136] ———, *The entropy formula for the Ricci flow and its geometric applications*, (2006), arXiv:math/0211159.

[137] F. Planchon and I. Rodnianski, *On uniqueness for the Cauchy problem in general relativity*, (2007), in preparation.

[138] D. Pollack, *Compactness results for complete metrics of constant positive scalar curvature on subdomains of S^n*, Indiana Univ. Math. Jour. **42** (1993), 1441–1456. MR MR1266101 (95c:53052)

[139] P.H. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, CBMS Regional Conference Series in Mathematics, vol. 65, 1986. MR 845785 (87j:58024)

[140] J. Ratzkin, *An end to end gluing construction for metrics of constant positive scalar curvature*, Indiana Univ. Math. Jour. **52** (2003), 703–726. MR MR1986894 (2004m:53066)

[141] A.D. Rendall, *Reduction of the characteristic initial value problem to the Cauchy problem and its applications to the Einstein equations*, Proc. Roy. Soc. London A **427** (1990), 221–239. MR MR1032984 (91a:83004)

[142] A.G. Riess *et al.*, *New Hubble Space Telescope discoveries of type Ia Supernovae at $z > 1$: Narrowing constraints on the early behavior of dark energy*, Astroph. Jour. **659** (2007), 98–121, arXiv:astro-ph/0611572.

[143] O. Sarbach, G. Calabrese, J. Pullin, and M. Tiglio, *Hyperbolicity of the BSSN system of Einstein evolution equations*, Phys. Rev. **D66** (2002), 064002, arXiv:gr-qc/0205064.

[144] R. Schoen, *Vacuum spacetimes which are identically Schwarzschild near spatial infinity*, talk given at the Santa Barbara Conference on Strong Gravitational Fields, June 22-26, 1999, `http://doug-pc.itp.ucsb.edu/online/gravity_c99/schoen/`.

[145] ———, *Conformal deformation of a Riemannian metric to constant scalar curvature*, J. Differential Geom. **20** (1984), 479–495. MR 788292 (86i:58137)

[146] R. Schoen and S.-T. Yau, *On the structure of manifolds with positive scalar curvature*, Manuscripta Math. **28** (1979), 159–183. MR MR535700 (80k:53064)

[147] ———, *Lectures on differential geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology, I, International Press, Cambridge, MA, 1994. MR MR1333601 (97d:53001)

[148] J. Sharples, *Local existence of quasispherical space-time initial data*, Jour. Math. Phys. **46** (2005), 052501, 28. MR 2142996 (2006m:83014)

[149] W.-X. Shi, *Ricci deformation of the metric on complete noncompact Riemannian manifolds*, Jour. Diff. Geom. **30** (1989), 303–394. MR 1010165 (90f:53080)

[150] M. Shibata and T. Nakamura, *Evolution of three-dimensional gravitational waves: harmonic slicing case*, Phys. Rev. D (3) **52** (1995), 5428–5444. MR MR1360417 (96g:83050)

[151] B. Smith and G. Weinstein, *On the connectedness of the space of initial data for the Einstein equations*, Electron. Res. Announc. Am. Math. Soc. **6** (2000), 52–63.

[152] H.F. Smith and D. Tataru, *Sharp local well-posedness results for the nonlinear wave equation*, Ann. of Math. (2) **162** (2005), 291–366. MR MR2178963 (2006k:35193)

[153] N.S. Trudinger, *Remarks concerning the conformal deformation of Riemannian structures on compact manifolds*, Ann. Scuola Norm. Sup. Pisa (3) **22** (1968), 265–274. MR 0240748 (39 #2093)

[154] J. Weidmann, *Linear operators in Hilbert spaces*, Graduate Texts in Mathematics, vol. 68, Springer Verlag, New York, Heidelberg, Berlin, 1980.

[155] W. Michael Wood-Vasey et al., *Observational Constraints on the Nature of the Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey*, Astroph. Jour. **666** (2007), 694–715, arXiv:astro-ph/0701041.

[156] H. Yamabe, *On a deformation of Riemannian structures on compact manifolds*, Osaka Math. J. **12** (1960), 21–37. MR 0125546 (23 #A2847)

[157] J.W. York, *Energy and momentum of the gravitational field*, Essays in General Relativity (A Festschrift for Abraham Taub) (F.J. Tipler, ed.), Academic Press, New York, 1980, pp. 39–58.