Dynamical and Hamiltonian formulation of General Relativity

Domenico Giulini

Institute for Theoretical Physics
Riemann Center for Geometry and Physics
Leibniz University Hannover, Appelstrasse 2, D-30167 Hannover, Germany
and
ZARM Bremen, Am Fallturm, D-28359 Bremen, Germany

Abstract

This is a substantially expanded version of a chapter-contribution to *The Springer Handbook of Spacetime*, edited by Abhay Ashtekar and Vesselin Petkov, published by Springer Verlag in 2014. It introduces the reader to the reformulation of Einstein's field equations of General Relativity as a constrained evolutionary system of Hamiltonian type and discusses some of its uses, together with some technical and conceptual aspects. Attempts were made to keep the presentation self contained and accessible to first-year graduate students. This implies a certain degree of explicitness and occasional reviews of background material.

Contents

1 Introduction

The purpose of this contribution is to explain how the field equations of General Relativity—often simply referred to as *Einstein's equations*—can be understood as dynamical system; more precisely, as a *constrained Hamiltonian system*.

In General Relativity, it is often said, spacetime becomes dynamical. This is meant to say that the geometric structure of spacetime is encoded in a field that, in turn, is subject to local laws of propagation and coupling, just as, e.g., the electromagnetic field. It is *not* meant to say that spacetime as a whole evolves. Spacetime does not evolve, spacetime just is. But a given spacetime (four dimensional) can be viewed as the evolution, or history, of space (three dimensional). There is a huge redundancy in this representation, in the sense that apparently very different evolutions of space represent the same spacetime. However, if the resulting spacetime is to satisfy Einstein's equations, the evolution of space must also obey certain well de-

fined restrictions. Hence the task is to give precise mathematical expression to the redundancies in representation as well as the restrictions of evolution for this picture of spacetime as space's history. This will be our main task.

This dynamical picture will be important for posing and solving time-dependent problems in General Relativity, like the scattering of black holes with its subsequent generation and radiation of gravitational waves. Quite generally, it is a key technology to

- formulate and solve initial value problems;

- integrate Einstein's equations by numerical codes;

- characterize dynamical degrees of freedom;

- characterize isolated systems and the association of asymptotic symmetry groups, which will give rise to globally conserved 'charges', like energy and linear as well as angular momentum (Poincaré charges).

Moreover, it is also the starting point for the *canonical quantization program*, which constitutes one main approach to the yet unsolved problem of *Quantum Gravity*. In this approach one tries to make essential use of the Hamiltonian structure of the classical theory in formulating the corresponding quantum theory. This strategy has been applied successfully in the transition from classical to quantum mechanics and also in the transition from classical to quantum electrodynamics. Hence the canonical approach to Quantum Gravity may be regarded as conservative, insofar as it tries to apply otherwise established rules to a classical theory that is experimentally and observationally extremely well tested. The underlying hypothesis here is that we may quantize interaction-wise. This distinguishes this approach from string theory, the underlying credo of which is that Quantum Gravity only makes sense on the basis of a unified descriptions of all interactions.

Historically the first paper to address the problem of how to put Einstein's equations into the form of a Hamiltonian dynamical system was Dirac's [54] from 1958. He also noticed its constrained nature and started to develop the corresponding generalization of *constrained Hamiltonian systems* in [53] and their quantization [55]. On the classical side, this developed into the more geometric *Dirac-Bergmann theory* of constraints [76] and on the quantum side into an elaborate theory of quantization of systems with gauge redundancies; see [80] for a comprehensive account.

Dirac's attempts were soon complemented by an extensive joint work of Richard Arnowitt, Stanley Deser, and Charles Misner - usually and henceforth abbreviated by ADM. Their work started in 1959 by a paper [3] of the first two of these authors and continued in the series [4] [6] [5] [9] [8] [7] [10] [11] [12] [14] [15] [13] of 12 more papers by all three. A comprehensive summary of their work was given in 1962 in [16], which was republished in 2008 in [17]; see also the editorial note [104] with short biographies of ADM.

A geometric discussion of Einstein's evolution equations in terms of infinite-dimensional symplectic geometry has been worked out by Fischer and Marsden in [57]; see also their beautiful summaries and extended discussions in [59] and [58]. More on the mathematical aspects of the initial-value problem, including the global behavior of gravitational fields in General Relativity, can be found in [39], [45], and [106]. Modern text-books on the 3+1 formalism and its application to physical problems and their numerical solution-techniques are [27, 77]. The Hamiltonian structure and its use in the canonical quantization program for gravity is discussed in [34, 92, 107, 113].

2 Notation and conventions

From now on "General Relativity" will be abbreviated by "GR". Spacetime is a differentiable manifold M of dimension n, endowed with a metric g of signature $(\varepsilon, +, \cdots, +)$. In GR $n = 4$ and $\varepsilon = -1$ and it is implicitly understood that these are the "right" values. However, either for the sake of generality and/or particular interest, we will sometimes state formulae for general n and

ε, where usually $n \geq 2$ (sometimes $n \geq 3$) and either $\varepsilon = -1$ (Lorentzian metric) or $\varepsilon = +1$ (Riemannian metric; also called Euclidean metric).

The case $\varepsilon = 1$ has been extensively considered in path-integral approaches to Quantum Gravity, then referred to as *Euclidean Quantum Gravity*.

The tangent space of M at point $p \in M$ will be denoted by T_pM, the cotangent space by T_p^*M, and the tensor product of u factors of T_pM with d factors of T_p^*M by $T_{pd}^{\,u}M$. (Mnemonic in components: u = number of indices "upstairs", d = number of indices "downstairs".) An element T in $T_{pd}^{\,u}M$ is called a tensor of contravariant rank u and covariant rank d at point p, or simply a tensors of rank (u,d) at p. T is called *contravariant* if $d = 0$ and $u > 0$, and *covariant* if $u = 0$ and $d > 0$. A tensor with $u > 0$ and $d > 0$ is then referred to as of *mixed type*. Note that $T_pM = T_{p0}^1M$ and $T_p^*M = T_{p1}^0M$. The set of tensor *fields*, i.e. smooth assignments of an element in $T_{pd}^{\,u}M$ for each $p \in M$, are denoted by $\Gamma T_d^u M$. Unless stated otherwise, *smooth* means C^∞, i.e. continuously differentiable to any order. For $t \in \Gamma T_d^u M$ we denote by $t_p \in T_{pd}^{\,u}M$ the evaluation of t at $p \in M$. $C^\infty(M)$ denotes the set of all C^∞ real-valued functions on M, which we often simply call smooth functions.

If $f : M \to N$ is a diffeomorphism between manifolds M and N, then $f_{*p} : T_pM \to T_{f(p)}N$ denotes the differential at p. The transposed (or dual) of the latter map is, as usual, denoted by $f_p^* : T_{f(p)}^*N \to T_p^*M$. If X is a vector field on M then f_*X is a vector field on N, called the *push forward* of X by f. It is defined by $(f_*X)_q := f_{*f^{-1}(q)}X_{f^{-1}(q)}$, for all $q \in N$. If α is a co-vector field on N then $f^*\alpha$ is a co-vector field on M, called the *pull back* of α by f. It is defined by $(f^*\alpha)_p := \alpha_{f(p)} \circ f_{*p}$, for all $p \in M$. For these definitions to make sense we see that f need generally not be a diffeomorphism; M and N need not even be of the same dimension. More precisely, if α is a smooth field of co-vectors that is defined at least on the image of f in N, then $f^*\alpha$, as defined above, is always a smooth field of co-vectors on M. However, for the push forward f_* of a general vector field on M to result in a well defined vector field on the image of f in N we certainly need injectivity

of f. If f is a diffeomorphism we can not only push-forward vectors and pull back co-vectors, but also vice versa. Indeed, if Y is a vector field on N one can write $f^*Y := (f^{-1})_*Y$ and call it the pull back of Y by the diffeomorphism f. Likewise, if β is a co-vector field on M, one can write $f_*\beta := (f^{-1})^*\beta$ and call it the push forward of β. In this fashion we can define both, push-forwards and pull-backs, of general tensor fields $T \in \Gamma T_d^u M$ by linearity and applying f_* or f^* tensor-factor wise.

Note that the general definition of *metric* is as follows: $g \in \Gamma T_2^0 M$, such that g_p is a symmetric non-degenerate bilinear form on T_pM. Such a metric provides isomorphisms (sometimes called the *musical isomorphisms*)

$$\flat : T_pM \to T_p^*M$$
$$X \mapsto X^\flat := g(X, \,\cdot\,)\,, \qquad (1a)$$
$$\sharp : T_p^*M \to T_pM$$
$$\omega \mapsto \omega^\sharp := \flat^{-1}(\omega)\,. \qquad (1b)$$

Using $\sharp$ we obtain a metric g_p^{-1} on T_p^*M from the metric g_p on T_pM as follows:

$$g_p^{-1}(\omega_1, \omega_2) := g_p(\omega_1^\sharp, \omega_2^\sharp) = \omega_1(\omega_2^\sharp)\,. \qquad (2)$$

We also recall that the tensor space $T_{p1}^1 M$ is naturally isomorphic to the linear space $\mathrm{End}(T_pM)$ of all endomorphisms (linear self maps) of T_pM. Hence it carries a natural structure as associative algebra, the product being composition of maps denoted by $\circ$. As usual, the *trace*, denoted Tr, and the *determinant*, denoted det, are the naturally defined real-valued functions on the space of endomorphisms. For purely co- or contravariant tensors the trace can be defined by first applying one of the isomorphisms (1). In this case we write Tr_g to indicate the dependence on the metric g.

Geometric representatives of curvature are often denoted by bold-faced abbreviations of their names, like **Riem** and **Weyl** for the (covariant, i.e. all indices down) Riemann and Weyl tensors, **Sec** for the sectional curvature, **Ric** and **Ein** for the Ricci and Einstein tensors, **Scal** for the scalar curvature, and **Wein** for the Weingarten map (which is essentially equivalent to the extrinsic curvature).

This is done in order to highlight the geometric meaning behind some basic formulae, at least the simpler ones. Later, as algebraic expressions become more involved, we will also employ the standard component notation for computational ease.

3 Einstein's equations

In n-dimensional spacetime Einstein's equations form a set of $\frac{1}{2}n(n+1)$ quasi-linear partial differential equations of second order for $\frac{1}{2}n(n+1)$ functions (the components of the metric tensor) depending on n independent variables (the coordinates in spacetime). At each point of spacetime (event) they equate a purely geometric quantity to the distribution of energy and momentum carried by the matter. More precisely, this distribution comprises the local densities (quantity per unit volume) and current densities (quantity per unit area and unit time) of energy and momentum. The geometric quantity in Einstein's equations is the Einstein tensor $\mathbf{Ein}$, the matter quantity is the energy-momentum tensor $\mathbf{T}$. Both tensors are of second rank, symmetric, and here taken to be covariant (in components: "all indices down"). Their number of independent components in n spacetime dimensions is $\frac{1}{2}n(n+1)$

Einstein's equations (actually a single tensor equation, but throughout we use the plural to emphasize that it comprises several component equations) state the simple proportionality of $\mathbf{Ein}$ with $\mathbf{T}$

$$\mathbf{Ein} = \kappa\,\mathbf{T}\,, \tag{3}$$

where κ denotes the dimensionful constant of proportionality. Note that no explicit reference to the dimension n of spacetime enters (3), so that even if $n \neq 4$ it is usually referred to as Einstein's equations. We could have explicitly added a *cosmological constant* term $g\Lambda$ on the left-hand side, where Λ is a constant the physical dimension of which is the square of an inverse length. However, as long as we write down our formulae for general $\mathbf{T}$ we may absorb this term into $\mathbf{T}$ where it accounts for a contribution $\mathbf{T}_\Lambda = -g\Lambda/\kappa$. This has to be kept

in mind when explicit models for $\mathbf{T}$ are used and when we speak of "vacuum", which now means:

$$\mathbf{T}_{\mathrm{vacuum}} = \mathbf{T}_\Lambda := -\kappa^{-1}g\Lambda\,. \tag{4}$$

The signs are chosen such that a positive Λ accounts for a positive energy density and a negative pressure if the spacetime is Lorentzian ($\varepsilon = -1$).

There is another form of Einstein's equations which is sometimes advantageous to use and in which n explicitly enters:

$$\mathbf{Ric} = \kappa\left(\mathbf{T} - \tfrac{1}{n-2}g\operatorname{Tr}_g(\mathbf{T})\right). \tag{5}$$

These two forms are easily seen to be mathematically equivalent by the identities

$$\mathbf{Ein} = \mathbf{Ric} - \tfrac{1}{2}g\operatorname{Tr}_g(\mathbf{Ric})\,, \tag{6a}$$
$$\mathbf{Ric} = \mathbf{Ein} - \tfrac{1}{n-2}g\operatorname{Tr}_g(\mathbf{Ein})\,. \tag{6b}$$

With respect to a local field of basis vectors $\{e_0, e_1, \cdots, e_{n-1}\}$ we write $\mathbf{Ein}(e_\mu, e_\nu) =: G_{\mu\nu}$, $\mathbf{T}(e_\mu, e_\nu) =: T_{\mu\nu}$, and $\mathbf{Ric}(e_\mu, e_\nu) =: R_{\mu\nu}$. Then (3) and (5) take on the component forms

$$G_{\mu\nu} = \kappa\,T_{\mu\nu} \tag{7}$$

and

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \tfrac{1}{n-2}g_{\mu\nu}T^\lambda_\lambda\right) \tag{8}$$

respectively. Next we explain the meanings of the symbols in Einstein's equations from left to right.

3.1 What aspects of geometry?

The left-hand side of Einstein's equations comprises certain measures of curvature. As will be explained in detail in Section 5, all curvature information in dimensions higher than two can be reduced to that of *sectional curvature*. The sectional curvature at a point $p \in M$ tangent to span$\{X, Y\} \subset T_pM$ is the Gaussian curvature at p of the submanifold spanned by the geodesics in M emanating from p tangent to span$\{X, Y\}$. The Gaussian curvature is defined as the product of two principal curvatures, each being measured in units

of an inverse length (the inverse of a principal radius). Hence the Gaussian curvature is measured in units of an inverse length-squared.

At each point p in spacetime the Einstein and Ricci tensors are symmetric bilinear forms on T_pM. Hence $\mathbf{Ein}_p$ and $\mathbf{Ric}_p$ are determined by the values $\mathbf{Ein}_p(W,W)$ and $\mathbf{Ric}_p(W,W)$ for all $W \in T_pM$. By continuity in W this remains true if we restrict W to the open and dense set of vectors which are not null, i.e. for which $g(W,W) \neq 0$. As we will see later on, we then have

$$\mathbf{Ein}(W,W) = -g(W,W)\sum_{\perp W}^{N_1}\mathbf{Sec}\,, \qquad (9)$$

$$\mathbf{Ric}(W,W) = +g(W,W)\sum_{\parallel W}^{N_2}\mathbf{Sec}\,. \qquad (10)$$

For the Einstein tensor the sum on the right-hand side is over any complete set of $N_1 = \frac{1}{2}(n-1)(n-2)$ sectional curvatures of pairwise orthogonal planes in the orthogonal complement of W in T_pM. For the Ricci tensor it is over any complete set of $N_2 = n-1$ sectional curvatures of pairwise orthogonal planes containing W. If W is a timelike unit vector representing an observer, $\mathbf{Ein}(W,W)$ is simply $(-\varepsilon)$ times an equally weighted sum of spacelike sectional curvatures, whereas $\mathbf{Ric}(W,W)$ is ε times an equally weighted sum of timelike sectional curvatures. In that sense we may say that, e.g., $\mathbf{Ein}(W,W)$ at $p \in M$ measures the mean Gaussian curvature of the (local) hypersurface in M that is spanned by geodesics emanating from W orthogonal to W. It, too, is measured in units of the square of an inverse length.

3.2   What aspects of matter?

Now we turn to the right-hand side of Einstein's equations. We restrict to four spacetime dimensions, though much of what we say will apply verbatim to other dimensions. The tensor $\mathbf{T}$ on the right-hand side of (3) is the *energy-momentum tensor* of matter. With respect to an orthonormal basis $\{e_0, e_1, \cdots, e_{n-1}\}$ with timelike e_0 the components $T_{\mu\nu} := \mathbf{T}(e_\mu, e_\nu)$ form a symmetric 4×4

matrix, which we represent as follows by splitting off terms involving a time component:

$$T_{\mu\nu} = \begin{pmatrix} \mathcal{E} & -c\vec{\mathcal{M}} \\ -\frac{1}{c}\vec{\mathcal{S}} & T_{mn} \end{pmatrix}\,. \qquad (11)$$

Here all matrix elements refer to the matter's energy momentum distribution relative to the rest frame of the observer who momentarily moves along e_0 (i.e. with four-velocity $u = ce_0$) and uses the basis $\{e_1, e_2, e_3\}$ in his/her rest frame. Then $\mathcal{E} = T_{00}$ is the energy density, $\vec{\mathcal{S}} = (s_1, s_2, s_3)$ the (components of the) energy current-density, i.e. energy per unit surface area and unit time interval, $\vec{\mathcal{M}}$ the momentum density, and finally T_{mn} the (component of the) momentum current-density, i.e. momentum per unit of area and unit time interval. Note that symmetry $T_{\mu\nu} = T_{\nu\mu}$ implies a simple relation between the energy current-density and the momentum density

$$\vec{\mathcal{S}} = c^2\,\vec{\mathcal{M}}\,. \qquad (12)$$

The remaining relations $T_{mn} = T_{nm}$ express equality of the m-th component of the current density for n-momentum with the n-th component of the current density for m-momentum. Note that the two minus signs in front of the mixed components of (11) would have disappeared had we written down the contravariant components $T^{\mu\nu}$. In flat spacetime, the four equations $\partial^\mu T_{\mu\nu}$ express the local conservation of energy and momentum. In curved spacetime (with vanishing torsion) we have the identity (to be proven later; compare (90b))

$$\nabla^\mu G_{\mu\nu} \equiv 0 \qquad (13)$$

implies via (7)

$$\nabla^\mu T_{\mu\nu} = 0\,, \qquad (14)$$

which may be interpreted as expressing a local conservation of energy and momentum for the matter *plus* the gravitational field, though there is no such thing as a separate energy-momentum tensor on spacetime for the gravitational field.

Several positivity conditions can be imposed on the energy momentum tensor $\mathbf{T}$. The simplest is known as *weak energy-condition* and reads

$\mathbf{T}(W, W) \geq 0$ for all timelike W. It is equivalent to the requirement that the energy density measured by any local observer is non negative. For a perfect fluid of rest-mass density ρ and pressure p the weak energy-condition is equivalent to both conditions $\rho \geq 0$ and $p \geq -c^2\rho$. The strong energy-condition says that $\left(\mathbf{T} - \frac{1}{2}g\mathrm{Tr}_g(\mathbf{T})\right)(W, W) \geq 0$ again for all timelike W. This neither follows nor implies the weak energy-condition. For a perfect fluid it is equivalent to both conditions $p \geq -c^2\rho$ and $p \geq -c^2\rho/3$, i.e. to the latter alone if ρ is positive and to the former alone if ρ is negative (which is not excluded here). Its significance lies in the fact that it ensures attractivity of gravity as described by Einstein's equations. It must, for example, be violated if matter is to drive inflation. Note that upon imposing Einstein's equations the weak and the strong energy-conditions read $\mathbf{Ein}(W, W) \geq 0$ and $\mathbf{Ric}(W, W) \geq 0$ respectively. From (9) and (10) we can see that for fixed W these imply conditions on complementary sets of sectional curvatures. For completeness we mention the *condition of energy dominance*, which states that $\mathbf{T}(W, W) \geq |\mathbf{T}(X, X)|$ for any pair of orthonormal vectors W, X where W is timelike (and hence X is spacelike). It is equivalent to the weak energy-condition supplemented by the requirement that $(i_W\mathbf{T})^\sharp$ be non spacelike for all timelike W. The second requirement ensures locally measured densities of energy currents and momenta of matter to be non spacelike.

3.3 How do geometry and matter relate quantitatively?

We return to Einstein's equations and finally discuss the constant of proportionality κ on the right-hand side of (3). Its physical dimension is that of curvature (m^{-2} in SI units) divided by that of energy density ($J \cdot m^{-3}$ in SI units, where $J =$ Joule). It is given by

$$\kappa := \frac{8\pi G}{c^4} \approx 2.1 \times 10^{-43}\, \frac{m^{-2}}{J \cdot m^{-3}}, \quad (15)$$

where $G \approx 6.67384(80) \times 10^{-11} \mathrm{m}^3 \cdot \mathrm{kg}^{-1} \cdot \mathrm{s}^{-2}$ is Newton's constant. It is currently (March 2013) known with a relative standard uncertainty of 1.2×10^{-4} and is thus by far the least well known of the fundamental physical constants. $c = 299\,792\,458\,\mathrm{m} \cdot \mathrm{s}^{-1}$ is the vacuum speed of light whose value is exact, due to the SI-definition of meter ("the meter is the length of the path traveled by light in vacuum during a time interval of $1/299\,792\,458$ of a second").

The physical dimension of κ is $\mathrm{time}^2/(\mathrm{mass} \cdot \mathrm{length})$, that is in SI-units $\mathrm{s}^2 \cdot \mathrm{kg}^{-1} \cdot \mathrm{m}^{-1}$ or $\mathrm{m}^{-2}/(\mathrm{J} \cdot \mathrm{m}^{-3})$, where $\mathrm{J} = \mathrm{Joule} = \mathrm{kg} \cdot \mathrm{m}^2 \cdot \mathrm{s}^{-2}$. It converts the common physical dimension of all components $T_{\mu\nu}$, which is that of an energy density (Joule per cubic meter in SI-units) into that of the components of $\mathbf{Ein}$, which is that of curvature (in dimension ≥ 2), i.e., the square of an inverse length (inverse square-meter in SI-units).

If we express energy density as mass density times c^2, the conversion factor is $\kappa c^2 = 8\pi G/c^2$. It can be expressed in various units that give a feel for the local "curving power" of mass-densities.

For that of water, $\rho_W \approx 10^3\,\mathrm{kg} \cdot \mathrm{m}^{-3}$, and nuclear matter in the core of a neutron star (which is more than twice that of atomic nuclei), $\rho_N \approx 5 \times 10^{17}\,\mathrm{kg} \cdot \mathrm{m}^{-3}$, we get, respectively:

$$\kappa c^2 \approx \left(\frac{1}{1.5\,\mathrm{AU}}\right)^2 \cdot \rho_W^{-1} \approx \left(\frac{1}{10\,\mathrm{km}}\right)^2 \cdot \rho_N^{-1}, \quad (16)$$

where $\mathrm{AU} = 1.5 \times 10^{11}\,\mathrm{m}$ is the astronomical unit (mean Earth-Sun distance). Hence, roughly speaking, matter densities of water cause curvature radii of the order of the astronomical unit, whereas the highest known densities of nuclear matter cause curvature radii of tens of kilometers. The curvature caused by mere mass density is that expressed in $\mathbf{Ein}(W, W)$ when W is taken to be the unit timelike vector characterizing the local rest frame of the matter: It is a mean of spatial sectional curvatures in the matter's local rest frame. Analogous interpretations can be given for the curvatures caused by momentum densities (energy current-densities) and momentum current-densities (stresses).

6

3.4 Conserved energy-momentum tensors and globally conserved quantities

In this subsection we briefly wish to point out that energy-momentum tensors $\mathbf{T}$ whose divergence vanishes (14) give rise to conserved quantities in case the spacetime (M, g) admits non-trivial isometries. We will stress the global nature of these quantities and clarify their mathematical habitat.

Conservation laws for the matter alone result in the presence of symmetries, more precisely, if *Killing fields* for (M, g) exist. Recall that a vector field V is called a Killing field iff $L_V g = 0$, where L_V is the Lie derivative with respect to V. Recall that the Lie derivative can be expressed in terms of the Levi-Civita covariant derivative with respect to g, in which case we get the component expression:

$$(L_V g)_{\mu\nu} = \nabla_\mu V_\nu + \nabla_\nu V_\mu = 0 \,. \qquad (17)$$

We consider the one-form J_V that results from contracting $\mathbf{T}$ with V:

$$J_V := i_V \mathbf{T} = V^\mu T_{\mu\nu} dx^\nu \,. \qquad (18)$$

As a result of Killing's equation (17) it is divergence free,

$$\nabla_\mu J_V^\mu = 0 \,. \qquad (19)$$

This may be equivalently expressed by saying that the 3-form $\star J_V$, which is the Hodge dual of the 1-form J_V, is closed:

$$d \star J_V = 0 \,. \qquad (20)$$

Integrating $\star J_V$ over some 3-dimensional submanifold Σ results in a quantity

$$Q[V, \Sigma] := \int_\Sigma \star J_V \qquad (21)$$

which, because of (20), is largely independent of Σ. More precisely, if $\Omega \subset M$ is an oriented domain with boundary $\partial\Omega = \Sigma_1 - \Sigma_2$, then Stokes' theorem gives $Q[V, \Sigma_1] = Q[V, \Sigma_2]$.

Suppose now that V arises from a finite-dimensional Lie group G that acts on (M, g) by isometries. We will discuss general Lie-group actions on manifolds in the Appendix at the end of this contribution, containing detailed proofs of some relevent formulae. But in order not to interrupt the argument too much, let us recall at this point that an *action* of G on M is a map

$$\begin{aligned} \Phi : G \times M &\to M \,, \\ (g, m) &\mapsto \Phi(g, m) = \Phi_g(m) \,, \end{aligned} \qquad (22a)$$

which satisfies

$$\Phi_e = \mathrm{Id}_M \,, \qquad (22b)$$
$$\Phi_g \circ \Phi_h = \Phi_{gh} \,. \qquad (22c)$$

Here $e \in G$ denotes the neutral element, Id_M the identity map on M, and equation (22c) is valid for any two elements g, h of G. In fact, equation (22c) characterizes a *left action*. In contrast, for a *right action* we would have Φ_{hg} instead of Φ_{gh} on the right-hand side of (22c). Moreover, as the group acts by isometries for the metric g, we also have $\Phi_h^* g = g$ for all $h \in G$.

Now, this action defines a map, V, from $\mathrm{Lie}(G)$, the Lie algebra of G, into the vector fields on M. The vector field corresponding to $X \in \mathrm{Lie}(G)$ is denoted V^X. Its value at a point $m \in M$ is defined by

$$V^X(m) := \frac{d}{dt}\Big|_{t=0} \Phi\big(\exp(tX), m\big) \,. \qquad (23)$$

From this it is obvious that $V : \mathrm{Lie}(G) \to \Gamma T_0^1 M$ is linear. Moreover, one may also show (compare (382b) in Appendix) that this map is a Lie anti-homomorphism, i.e. that

$$V^{[X,Y]} = -[V^X, V^Y] \,. \qquad (24)$$

(As shown in the Appendix, a right action would have resulted in a proper Lie homomorphism – see (382a) –, i.e. without the minus sign on the right-hand side, which however is not harmful.) The left action of G on M extends to a left action on all tensor fields by push forward. In particular, the push forward of V^X by Φ_g has a simple expression (see (383a) in Appendix) :

$$\Phi_{g*} V^X = V^{\mathrm{Ad}_g(X)} \,, \qquad (25)$$

where Ad denotes the adjoint representation of G on $\mathrm{Lie}(G)$. In fact, relation (25) can be directly deduced from definition (23). Indeed, writing $\Phi(g,p) = g \cdot p$ for notational simplicity, we have (see Appendix for more explanation)

$$
\begin{aligned}
(\Phi_{g*} V^X)(g \cdot p) &= \frac{d}{dt}\Big|_{t=0}\big(g \exp(tX) \cdot p\big) \\
&= \frac{d}{dt}\Big|_{t=0}\big(g \exp(tX)\, g^{-1} g \cdot p\big) \\
&= \frac{d}{dt}\Big|_{t=0}\Big(\exp\big(t\mathrm{Ad}_g(X)\big) \cdot (g \cdot p)\Big) \\
&= V^{\mathrm{Ad}_g(X)}(g \cdot p)\,.
\end{aligned}
\tag{26}
$$

This leads to (25) which we shall use shortly.

Returning to the expression (21) we see that, for fixed Σ, it becomes a linear map from $\mathrm{Lie}(G)$ to $\mathbb{R}$:

$$
\mathfrak{M} : \mathrm{Lie}(G) \to \mathbb{R}, \quad \mathfrak{M}(X) := Q[V^X, \Sigma]. \tag{27}
$$

Hence each hypersurface Σ defines an element $\mathfrak{M} \in \mathrm{Lie}^*(G)$ in the vector space that is dual to the Lie algebra, given that the integral over Σ converges. This is the case for spacelike Σ and energy momentum tensors with spatially compact support (or at least sufficiently rapid fall off). The same argument as above using Stokes' theorem and (20) then shows that $\mathfrak{M}$ is independent of the choice of spacelike Σ. In other words, we obtain a conserved quantity $\mathfrak{M} \in \mathrm{Lie}^*(G)$ for G-symmetric spacetimes (M,g) and covariant divergence free tensors $\mathbf{T}$.

So far we considered a fixed spacetime (M,g) and a fixed energy-momentum tensor $\mathbf{T}$, both linked by Einstein's equations. In this case the vanishing divergence (14) is an integrability condition for Einstein's equation and hence automatic. However, it is also of interest to consider the more general case where (M,g) is merely a background for some matter represented by energy-momentum tensors $\mathbf{T}[\mu]$, all of which are divergence free (14) with respect to the background metric g. Note that we do not assume (M,g) to satisfy Einstein's equations with any of the $\mathbf{T}[\mu]$ on the right-hand side. The μ stands for some matter variables which may be fundamental fields and/or of phenomenological

nature. In any case, we assume the isometric action (22) to extend to an action of G on the set of matter variables μ, which we denote by $\mu \mapsto \Phi_{g*}\mu$, like the push-forward on tensor fields. This is also meant to indicate that we assume this to be a left action, i.e. $\Phi_{g*} \circ \Phi_{h*} = \Phi_{gh*}$.

We regard the energy-momentum tensor $\mathbf{T}$ as a map from the space of matter variables to the space of symmetric second-rank covariant tensor fields on M. We require this map to satisfy the following covariance property:

$$
\mathbf{T}[\Phi_{g*}\mu] = \Phi_{g*}\mathbf{T}[\mu] := \Phi_{g^{-1}}^* \mathbf{T}\,, \tag{28}
$$

where Φ_{g*} is the ordinary push-forward of the tensor $\mathbf{T}$. Since we take $\mathbf{T}$ to be covariant, its push forward is the pull back by the inverse diffeomorphism, as indicated by the second equality in (28).

For each specification μ of matter variables we can compute the quantitty $Q[V^X, \Sigma, \mu]$ as in (21). Note that we now indicate the dependence on μ explicitly. We are interested in computing how Q changes as μ is acted on by $g \in G$. This is done as follws:

$$
\begin{aligned}
Q\big[V^X, \Sigma, \Phi_{g*}\mu\big] &= \int_\Sigma \star\, i_{V^X}\, \mathbf{T}[\Phi_{g*}\mu] \\
&= \int_\Sigma \star\, i_{V^X}\, \Phi_{g^{-1}}^* \mathbf{T}[\mu] \\
&= \int_\Sigma \star\, \Phi_{g^{-1}}^*\big(i_{\Phi_{g^{-1}*} V^X}\, \mathbf{T}[\mu]\big) \\
&= \int_\Sigma \star\, \Phi_{g^{-1}}^*\big(i_{V^{\mathrm{Ad}_{g^{-1}}(X)}}\, \mathbf{T}[\mu]\big) \\
&= \int_\Sigma \Phi_{g^{-1}}^*\big(\star\, i_{V^{\mathrm{Ad}_{g^{-1}}(X)}}\, \mathbf{T}[\mu]\big) \\
&= \int_{\Phi_{g^{-1}}(\Sigma)} \star\, i_{V^{\mathrm{Ad}_{g^{-1}}(X)}}\, \mathbf{T}[\mu] \\
&= Q\big[V^{\mathrm{Ad}_{g^{-1}}(X)}, \Phi_{g^{-1}}(\Sigma), \mu\big]\,.
\end{aligned}
\tag{29}
$$

Here we used (28) in the second equality, the general formula $i_V f^* T = f^*(i_{f_* V} T)$ (valid for any diffeomorphism f, vector field V, and covariant tensor field T) in the third equality, (25) in the fourth equality, the formula $\star f^* F = f^* \star F$ in the fifth

equality (valid for any orientation preserving isometry f and any form-field F; here we assume M to be oriented), and finally the general formula for the integral of the pull back of a form in the sixth equality.

Our final assumption is that Q does not depend on which hypersurface $\Phi_g(\Sigma)$ it is evaluated on. Since we assume (14) this is guaranteed if all $\Phi_g(\Sigma)$ are in the same homology class or, more generally, if any two hypersurfaces Σ and $\Phi_g(\Sigma)$ are homologous to hypersurfaces in the complement of the support of $\mathbf{T}$. A typical situation arising in physical applications is that of a source $\mathbf{T}[\mu]$ with spatially compact support; then any two sufficiently extended spacelike slices through the timelike support-tube of $\mathbf{T}[\mu]$ is homologous to the timelike cylindrical hypersurface outside this support-tube. In this case we infer from (29) that

$$Q\big[V^X, \Sigma, \Phi_{g*}\mu\big]e = Q\big[V^{\mathrm{Ad}_{g^{-1}}(X)}, \Sigma, \mu\big]. \quad (30)$$

Recall from (27) that for fixed Σ and $\mathbf{T}$ we have $\mathfrak{M} \in \mathrm{Lie}^*(G)$. Given the independence on Σ and the depencence of $\mathbf{T}$ on μ, we now regard $\mathfrak{M}$ as a map from the matter variables μ to $\mathrm{Lie}^*(G)$. This map may be called the *momentum map*. (Compare the notion of a momentum map in Hamiltonian mechanics; cf. Section 7.) Equation (30) then just states the Ad^*-equivariance of the momentum map:

$$\mathfrak{M} \circ \Phi_{g*} = \mathrm{Ad}_g^* \circ \mathfrak{M}. \quad (31)$$

Here Ad^* denote the co-adjoint representation of G on $\mathrm{Lie}(G)$, which is defined by $\mathrm{Ad}_g^*(\alpha) = \alpha \circ \mathrm{Ad}_{g^{-1}}$. From all this we see that the conserved "momentum" that we obtain by evaluating $\mathfrak{M}$ on the matter configuration μ is a conserved quantity that is *globally* associated to all of spacetime, not a particular region or point of it. It is an element of the vector space $\mathrm{Lie}^*(G)$ which carries the co-adjoint representation of the symmetry group G.

In particular this applies to Special Relativity, where M is the four-dimensional real affine space with associated (four-dimensional real) vector space V and g a bilinear, symmetric, non-degenerate form of signature $(-, +, +, +)$ [the sig-

nature does not matter in what follows]. The linear isometries of (V, η) form the Lorentz group $Lor \subset GL(V)$ and the isometries G of (M, g) can be (non-naturally) identified with the semi-direct product $V \rtimes Lor$, called the Poincaré group, $Poin$. Using g we can identify $\mathrm{Lie}^*(Poin)$ with $V \oplus (V \wedge V)$. The co-adjoint action of $(a, A) \in Poin$ on $(f, F) \in \mathrm{Lie}^*(Poin)$ is then given by

$$\mathrm{Ad}^*_{(a,A)}(f, F) = \Big(Af, \, (A \otimes A)F - a \wedge Af\Big). \quad (32)$$

Note that, e.g., the last term on the right hand side includes the law of change of angular momentum under spatial translations. In contrast, the adjoint representation on $\mathrm{Lie}(Poin)$, the latter also identified with $V \oplus (V \wedge V)$, is given by

$$\mathrm{Ad}_{(a,A)}(f, F) = \Big(Af - \big((A \otimes A)F\big)a, \, (A \otimes A)F\Big), \quad (33)$$

where the application of an element in $V \wedge V$ to an element in V is given by $(u \wedge v)(w) := u\,g(v, w) - v\,g(u, w)$, and linear extension. Note the characteristic difference between (32) and (33), which lies in the different actions of the subgroup of translations, whereas the subgroup of Lorentz transformations acts in the same fashion. Physical momenta transform as in (32), as already exemplified by the non-trivial transformation behavior of angular momentum under spatial translations. For a detailed discussion of the proper group-theoretic setting and the adjoint and co-adjoint actions, see the recent account [75].

4 Spacetime decomposition

In this section we explain how to decompose a given spacetime (M, g) into "space" and 'time". For this to be possible we need to make the assumption that M is diffeomorphic to the product of the real line $\mathbb{R}$ and some 3-manifold Σ:

$$M \cong \mathbb{R} \times \Sigma. \quad (34)$$

This will necessarily be the case for *globally hyperbolic* spacetimes, i.e. spacetimes admitting a

Cauchy surface [65]. We assume Σ to be orientable, for, if it were not, we could take the orientable double cover of it instead. Orientable 3-manifolds are always parallelizable [111] , i.e. admit three globally defined and pointwise linearly independent vector fields. This is equivalent to the triviality of the tangent bundle. In the closed case this is known as *Stiefel's theorem* (compare [100], problem 12-B) and in the open case it follows, e.g., from the well known fact that every open 3-manifold can be immersed in $\mathbb{R}^3$ [117]. Note that orientability is truly necessary; e.g., $\mathbb{R}P^2 \times S^1$ is not parallelizable. Since Cartesian products of parallelizable manifolds are again parallelizable, it follows that a 4-dimensional product spacetime (34) is also parallelizable. This does, of course, not generalize to higher dimensions. Now, for non-compact four-dimensional spacetimes it is known from [64] that parallelizability is equivalent to the existence of a *spin structure*, without which spinor fields could not be defined on spacetime. So we see that the existence of spin structure is already implied by (34) and hence does not pose any further topological restriction. Note that the only other potential topological restriction at this stage is that imposed from the requirement that a smooth Lorentz metric is to exist everywhere on spacetime. This is equivalent to a vanishing Euler characteristic (see, e.g., § 40 in [111]) which in turn is equivalently to the global existence of a continuous, nowhere vanishing vector field (possibly up to sign) on spacetime. But such a vector field clearly exists on any Cartesian product with one factor being $\mathbb{R}$. We conclude that existence of a Lorentz metric and a spin structure on an orientable spacetime $M = \mathbb{R} \times \Sigma$ pose no restrictions on the topology of an orientable Σ. As we will see later on, even Einstein's equation poses *no* topological restriction on Σ, in the sense that *some* (physically reasonable) solutions to Einstein's equations exist for any given Σ. Topological restrictions may occur, however, if we ask for solution with special properties (see below).

Now, given Σ, we consider a one-parameter family of embeddings

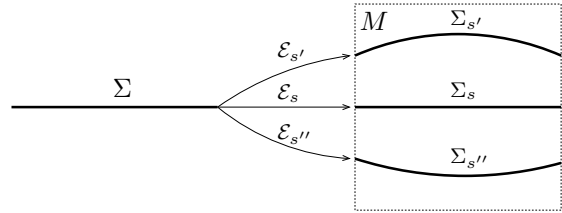$$\mathcal{E}_s : \Sigma \to M \,, \quad \Sigma_s := \mathcal{E}_s(\Sigma) \subset M \,. \qquad (35)$$

Figure 1: Spacetime M is foliated by a one-parameter family of spacelike embeddings of the 3-manifold Σ. Here the image $\Sigma_{s'}$ of Σ under $\mathcal{E}_{s'}$ lies to the future (above) and $\Sigma_{s''}$ to the past (below) of Σ_s if $s'' < s < s'$. 'Future' and 'past' refer to the time function t which has so far not been given any metric significance.

We distinguish between the abstract 3-manifold Σ and its image Σ_s in M. The latter is called the leaf corresponding to the value $s \in \mathbb{R}$. Each point in M is contained in precisely one leaf. Hence there is a real valued function $t : M \to \mathbb{R}$ that assigns to each point in M the parameter value of the leaf it lies on:

$$t(p) = s \Leftrightarrow p \in \Sigma_s \,. \qquad (36)$$

So far this is only a foliation of spacetime by 3-dimensional leaves. For them to be addressed as "space" the metric induced on them must be positive definite, that is, the leaves should be spacelike submanifolds. This means that the one-form dt is timelike:

$$g^{-1}(dt, dt) < 0 \,. \qquad (37)$$

The normalized field of one-forms is then

$$n^\flat := \frac{dt}{\sqrt{-g^{-1}(dt, dt)}} \,. \qquad (38)$$

As explained in section 2, we write $n^\flat$ since we think of this one form as the image under g of the normalized vector field perpendicular to the leaves:

$$n^\flat = g(n, \cdot) \,. \qquad (39)$$

The linear subspace of vectors in T_pM which are tangent to the leaf through p is denoted by $T_p^{\parallel}M$; hence

$$T_p^{\parallel}M := \{X \in T_pM : dt(X) = 0\} \,. \qquad (40)$$

The orthogonal complement is just the span of n at p, which we denote by $T_p^\perp M$. This gives, at each point p of M, the g-orthogonal direct sum

$$T_p M = T_p^\perp M \oplus T_p^\parallel M \, . \qquad (41)$$

and associated projections (we drop reference to the point p)

$$\begin{aligned} P^\perp : TM &\to T^\perp M \, , \\ X &\mapsto \varepsilon\, g(X, n)\, n \, , \qquad (42a) \end{aligned}$$

$$\begin{aligned} P^\parallel : TM &\to T^\parallel M \, , \\ X &\mapsto X - \varepsilon g(X, n)\, n \, . \qquad (42b) \end{aligned}$$

As already announced in Section 2, we introduced the symbol

$$\varepsilon = g(n, n) \qquad (43)$$

in order to keep track of where the signature matters. Note that the projection operators (42) are self-adjoint with respect to g, so that for all $X, Y \in TM$ we have

$$g\big(P^\perp X, Y\big) = g\big(X, P^\perp Y\big) \, , \qquad (44a)$$

$$g\big(P^\parallel X, Y\big) = g\big(X, P^\parallel Y\big) \, . \qquad (44b)$$

A vector is called *horizontal* iff it is in the kernel of $P^\perp$, which is equivalent to being invariant under $P^\parallel$. It is called *vertical* iff it is in the kernel of $P^\parallel$, which is equivalent to being invariant under $P^\perp$.

All this can be extended to forms. We define *vertical and horizontal forms* as those annihilating horizontal and vertical vectors, respectively:

$$T_p^{*\perp} M := \{\omega \in T_p^* M : \omega(X) = 0 \, , \ \forall X \in T_p^\parallel M\} \, , \qquad (45a)$$

$$T_p^{*\parallel} M := \{\omega \in T_p^* M : \omega(X) = 0 \, , \ \forall X \in T_p^\perp M\} \, . \qquad (45b)$$

Using the 'musical' isomorphisms (1), the self-adjoint projection maps (42) on vectors define self-adjoint projection maps on co-vectors (again dropping the reference to the base-point p)

$$P_*^\perp := \flat \circ P^\perp \circ \sharp \ : \ T^* M \to T^{*\perp} M \, , \qquad (46a)$$

$$P_*^\parallel := \flat \circ P^\parallel \circ \sharp \ : \ T^* M \to T^{*\parallel} M \, . \qquad (46b)$$

For example, letting the horizontal projection of the form ω act on the vector X, we get

$$\begin{aligned} P_*^\parallel \omega(X) &= (P^\parallel \omega^\sharp)^\flat(X) \\ &= g\big(P^\parallel \omega^\sharp, X\big) \\ &= g\big(\omega^\sharp, P^\parallel X\big) \\ &= \omega\big(P^\parallel X\big) \, , \end{aligned} \qquad (47)$$

where we merely used the definitions (1) of $\flat$ and $\sharp$ in the second and fourth equality, respectively, and the self-adjointness (44b) of $P^\parallel$ in the third equality. The analogous relation holds for $P_*^\perp \omega(X)$. It is also straightforward to check that $P_*^\parallel$ and $P_*^\perp$ are self-adjoint with respect to g^{-1} (cf. (2)).

Having the projections defined for vectors and co-vectors, we can also define it for the whole tensor algebra of the underlying vector space, just by taking the appropriate tensor products of these maps. All tensor products between $P^\parallel$ and $P_*^\parallel$ will then, for simplicity, just be denoted by $P^\parallel$, the action on the tensor being obvious. Similarly for $P^\perp$. (For what follows we need not consider mixed projections.) The projections being pointwise operations, we can now define vertical and horizontal projections of arbitrary tensor fields. Hence a tensor field $T \in \Gamma T_d^u M$ is called *horizontal* if and only if $P^\parallel T = T$. The space of horizontal tensor fields of rank (u, d) is denoted by $\Gamma T^{\parallel u}_d M$.

As an example, the horizontal projection of the metric g is

$$h := P^\parallel g := g\big(P^\parallel \cdot, P^\parallel \cdot\big) = g - \varepsilon n^\flat \otimes n^\flat \, . \qquad (48)$$

Hence $h \in \Gamma T^{\parallel 0}_2 M$. Another example of a horizontal vector field is the "acceleration" of the normal field n:

$$a := \nabla_n n \, . \qquad (49)$$

Here ∇ denotes the Levi-Civita covariant derivative with respect to g. An observer who moves perpendicular to the horizontal leaves has four-velocity $u = cn$ and four-acceleration $c^2 a$. If L denotes the Lie derivative, it is easy to show that the acceleration 1-form satisfies

$$a^\flat = L_n n^\flat \, . \qquad (50)$$

Moreover, as n is hypersurface orthogonal it is irrotational, hence its 1-form equivalent satisfies

$$dn^\flat \wedge n^\flat = 0 \,, \tag{51a}$$

which is equivalent to the condition of vanishing horizontal curl:

$$P^\parallel dn^\flat = 0 \,. \tag{51b}$$

Equation (51a) can also be immediately inferred directly from (38). Taking the operation $i_n \circ d$ (exterior derivative followed by contraction with n) as well as the Lie derivative with respect to n of (50) shows

$$da^\flat \wedge n^\flat = 0 \,, \tag{52a}$$

an equivalent expression being again the vanishing of the horizontal curl of a:

$$P^\parallel da^\flat = 0 \,. \tag{52b}$$

This will be useful later on.

Note that $a^\flat$ is a horizontal co-vector field, i.e. an element of $\Gamma T^{\parallel u=0}_{d=1} M$. More generally, for a purely covariant horizontal tensor field we have the following results, which will also be useful later on: Let $T \in \Gamma T^{\parallel 0}_d M$, then

$$P^\parallel L_n T = L_n T \,, \tag{53a}$$
$$L_{fn} T = f L_n T \,, \tag{53b}$$

for all $f \in C^\infty(M)$. Note that (53a) states that the Lie derivative in normal direction of a horizontal covariant tensor field is again horizontal. That this is not entirely evident follows, e.g., from the fact that a corresponding result does not hold for $T \in \Gamma T^{\parallel u}_d M$ where $u > 0$. The proofs of (53) just use standard manipulations.

A fixed space-point $q \in \Sigma$ defines the worldline (history of that point) $\mathbb{R} \ni s \mapsto \mathcal{E}_s(q)$. The collection of all worldlines of all space-points define a foliation of M into one-dimensional timelike leafs. Each leaf is now labeled uniquely by a space point. We can think of "space", i.e., the abstract manifold Σ, as the quotient $M/\sim$, where $p \sim p'$ iff both points lie on the same worldline. As any Σ_s intersects each worldline exactly once, each Σ_s is a

representative of space. Instead of using the foliation by 3-dimensional spatial leaves (35) we could have started with a foliation by timelike lines, plus the condition that these lines are vorticity free. These two concepts are equivalent. Depending on the context, one might prefer to emphasize one or the other.

The vector parallel to the worldline at $p = \mathcal{E}_s(q)$ is, as usual in differential geometry, defined by its action on $f \in C^\infty(M)$ (smooth, real valued functions):

$$\left.\frac{\partial}{\partial t}\right|_{\mathcal{E}_s(q)} f = \left.\frac{df(\mathcal{E}_{s'}(q))}{ds'}\right|_{s'=s} \,. \tag{54}$$

At each point this vector field can be decomposed into its horizontal component that is tangential to the leaves of the given foliation and its normal component. We write

$$\frac{1}{c}\frac{\partial}{\partial t} = \alpha\, n + \beta \,, \tag{55}$$

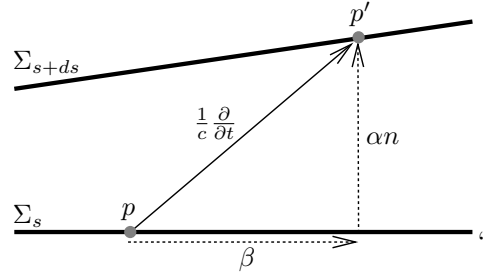where β is the tangential part; see Figure 2. The

Figure 2: For fixed $q \in \Sigma$ its image points $p = \mathcal{E}_s(q)$ and $p' = \mathcal{E}_{s+ds}(q)$ for infinitesimal ds are connected by the vector $\partial/\partial t|_p$, whose components normal to Σ_s are α (one function, called lapse) and β (three functions, called shift) respectively.

real-valued function α is called the *lapse* (function) and the horizontal vector field β is called the *shift* (vector-field) .

4.1 Decomposition of the metric

Let $\{e_0, e_1, e_2, e_3\}$ be a locally defined orthonormal frame with dual frame $\{\theta^0, \theta^1, \theta^2, \theta^3\}$. We call them *adapted* to the foliation if $e_0 = n$ and $\theta^0 = n^\flat$.

A local coordinate system $\{x^0, x^1, x^2, x^3\}$ is called *adapted* if $\partial/\partial x^a$ are horizontal for $a = 1, 2, 3$. Note that in the latter case $\partial/\partial x^0$ is not required to be orthogonal to the leaves (i.e. it need not be parallel to n). For example, we may take x^0 to be proportional to t; say $x^0 = ct$.

In the orthonormal co-frame the spacetime metric, i.e. the field of signature $(\varepsilon, +, +, +)$ metrics in the tangent spaces, has the simple form

$$g = \varepsilon \theta^0 \otimes \theta^0 + \sum_{a=1}^{3} \theta^a \otimes \theta^a. \qquad (56)$$

The inverse spacetime metric, i.e. the field of signature $(\varepsilon, +, +, +)$ metrics in the co-tangent spaces, has the form

$$g^{-1} = \varepsilon e_0 \otimes e_0 + \sum_{a=1}^{3} e_a \otimes e_a. \qquad (57)$$

The relation that expresses the coordinate basis in terms of the orthonormal basis is of the form (in a self-explanatory matrix notation)

$$\begin{pmatrix} \partial/\partial x^0 \\ \partial/\partial x^m \end{pmatrix} = \begin{pmatrix} \alpha & \beta^a \\ 0 & A^a_m \end{pmatrix} \begin{pmatrix} e_0 \\ e_a \end{pmatrix}, \qquad (58)$$

where β^a are the components of β with respect to the horizontal frame basis $\{e_a\}$. The inverse of (58) is

$$\begin{pmatrix} e_0 \\ e_a \end{pmatrix} = \begin{pmatrix} \alpha^{-1} & -\alpha^{-1}\beta^m \\ 0 & [A^{-1}]^m_a \end{pmatrix} \begin{pmatrix} \partial/\partial x^0 \\ \partial/\partial x^m \end{pmatrix}, \qquad (59)$$

where β^m are the components of β with respect to the horizontal coordinate-induced frame basis $\{\partial/\partial x^m\}$.

The relation for the co-bases dual to those in (58) is given by the transposed of (58), which we write as:

$$\begin{pmatrix} \theta^0 & \theta^a \end{pmatrix} = \begin{pmatrix} dx^0 & dx^m \end{pmatrix} \begin{pmatrix} \alpha & \beta^a \\ 0 & A^a_m \end{pmatrix}. \qquad (60)$$

The inverse of that is the transposed of (59):

$$\begin{pmatrix} dx^0 & dx^m \end{pmatrix} = \begin{pmatrix} \theta^0 & \theta^a \end{pmatrix} \begin{pmatrix} \alpha^{-1} & -\alpha^{-1}\beta^m \\ 0 & [A^{-1}]^m_a \end{pmatrix}. \qquad (61)$$

Orthogonality of the e_a implies for the chart components of the spatial metric (48)

$$h_{mn} := h(\partial/\partial x^m, \partial/\partial x^n) = \sum_{a=1}^{3} A^a_m A^a_n, \qquad (62)$$

and its inverse

$$h^{mn} := h^{-1}(dx^m, dx^n) = \sum_{a=1}^{3} [A^{-1}]^m_a [A^{-1}]^n_a. \qquad (63)$$

Inserting (60) into (56) and using (62) leads to the (3+1)-form of the metric in adapted coordinates

$$\begin{aligned} g &= \left(\varepsilon \alpha^2 + h(\beta, \beta)\right) c^2 \, dt \otimes dt \\ &\quad + c\beta_m \left(dt \otimes dx^m + dx^m \otimes dt\right) \\ &\quad + h_{mn} \, dx^m \otimes dx^n, \end{aligned} \qquad (64)$$

where $\beta_m := h_{mn}\beta^n$ are the components of $\beta^\flat := g(\beta, \cdot) = h(\beta, \cdot)$ with respect to the coordinate basis $\{\partial/\partial x^m\}$. Likewise, inserting (61) into (57) and using (63) leads to the (3+1)-form of the inverse metric in adapted coordinates (we write $\partial_t := \partial/\partial t$ and $\partial_m := \partial/\partial x^m$ for convenience)

$$\begin{aligned} g^{-1} &= \varepsilon c^{-2} \alpha^{-2} \, \partial_t \otimes \partial_t \\ &\quad - \varepsilon c^{-1} \alpha^{-2} \, \beta^m \left(\partial_t \otimes \partial_m + \partial_m \otimes \partial_t\right) \\ &\quad + \left(h^{mn} + \varepsilon \beta^m \beta^n\right) \partial_m \otimes \partial_n. \end{aligned} \qquad (65)$$

Finally we note that the volume form on spacetime also easily follows from (60)

$$\begin{aligned} d\mu_g &= \theta^0 \wedge \theta^1 \wedge \theta^2 \wedge \theta^3 \\ &= \alpha \sqrt{\det\{h_{mn}\}} \, cdt \wedge d^3x, \end{aligned} \qquad (66)$$

where we use the standard shorthand $d^3x = dx^1 \wedge dx^2 \wedge dx^3$.

4.2 Decomposition of the covariant derivative

Given horizontal vector fields X and Y, the covariant derivative of Y with respect to X need not be horizontal. Its decomposition is written as

$$\nabla_X Y = D_X Y + nK(X, Y), \qquad (67)$$

13

where

$$D_X Y := P^{\parallel} \nabla_X Y \,, \qquad (68)$$

$$K(X, Y) := \varepsilon \, g(n, \nabla_X Y) \,. \qquad (69)$$

The map D defines a covariant derivative (in the sense of Kozul; compare [110], Vol 2) for horizontal vector fields, as a trivial check of the axioms reveals. Moreover, since the commutator $[X, Y]$ of two horizontal vector fields is always horizontal (since the horizontal distribution is integrable by construction), we have

$$\begin{aligned}
T^D(X, Y) &= D_X Y - D_Y X - [X, Y] \\
&= P^{\parallel}\big(\nabla_X Y - \nabla_Y X - [X, Y]\big) \quad (70) \\
&= 0
\end{aligned}$$

due to ∇ being torsion free. We recall that torsion is a tensor field $T \in \Gamma T_2^1 M$ associated to each covariant derivative ∇ via

$$T^{\nabla}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \,. \qquad (71)$$

We have $T(X, Y) = -T(Y, X)$. As usual, even though the operations on the right hand side of (71) involve tensor *fields* (we need to differentiate), the result of the operation just depends on X and Y pointwise. This one proves by simply checking the validity of $T(fX, Y) = fT(X, Y)$ for all smooth functions f. Hence (70) shows that D is torsion free because ∇ is torsion free.

Finally, we can uniquely extend D to all horizontal tensor fields by requiring the Leibniz rule. Then, for X, Y, Z horizontal

$$\begin{aligned}
(D_X &h)(Y, Z) \\
&= X\big(h(Y, Z)\big) - h(D_X Y, Z) - h(Y, D_X Z) \\
&= X\big(g(Y, Z)\big) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) \\
&= (\nabla_X g)(Y, Z) = 0
\end{aligned} \qquad (72)$$

due to the metricity, $\nabla g = 0$, of ∇. Hence D is metric in the sense

$$Dh = 0 \,. \qquad (73)$$

The map K from pairs of horizontal vector fields (X, Y) into functions define a symmetric tensor field. Symmetry follows from the vanishing torsion of ∇, since then

$$\begin{aligned}
K(X, Y) &= \varepsilon \, g(n, \nabla_X Y) \\
&= \varepsilon \, g(n, \nabla_Y X + [X, Y]) \\
&= \varepsilon \, g(n, \nabla_Y X) \\
&= K(Y, X)
\end{aligned} \qquad (74)$$

for horizontal X, Y. From (69) one sees that $K(fX, Y) = fK(X, Y)$ for any smooth function f. Hence K defines a unique symmetric tensor field on M by stipulating that it be horizontal, i.e. $K(n, \cdot) = 0$. It is called the *extrinsic curvature* of the foliation or *second fundamental form*, the first fundamental form being the metric. From (69) and the symmetry just shown one immediately infers the alternative expressions

$$K(X, Y) = -\varepsilon \, g(\nabla_X n, Y) = -\varepsilon \, g(\nabla_Y n, X) \,. \qquad (75)$$

This shows the relation between the extrinsic curvature and the *Weingarten map*, **Wein**, also called the *shape operator*, which sends horizontal vectors to horizontal vectors according to

$$X \mapsto \mathbf{Wein}(X) := \nabla_X n \,. \qquad (76)$$

Horizontality of $\nabla_X n$ immediately follows from n being normalized: $g(n, \nabla_X n) = \frac{1}{2} X\big(g(n, n)\big) = 0$. Hence (75) simply becomes

$$\begin{aligned}
K(X, Y) &= -\varepsilon \, h\big(\mathbf{Wein}(X), Y\big) \\
&= -\varepsilon \, h\big(X, \mathbf{Wein}(Y)\big) \,,
\end{aligned} \qquad (77)$$

where we replaced g with h—defined in (48)—since both entries are horizontal. It says that K is $(-\varepsilon)$ times the covariant tensor corresponding to the Weingarten map, and that the symmetry of K is equivalent to the self-adjointness of the Weingarten map with respect to h. The Weingarten map characterizes the bending of the embedded hypersurface in the ambient space by answering the following question: In what direction and by what amount does the normal to the hypersurface tilt if, starting at point p, you progress within the hypersurface by the vector X. The answer is just $\mathbf{Wein}_p(X)$. Self adjointness of **Wein**

then means that there always exist three ($n-1$ in general) perpendicular directions in the hypersurface along which the normal tilts in the same direction. These are the *principal curvature directions* mentioned above. The principal curvatures are the corresponding eigenvalues of **Wein**.

Finally we note that the covariant derivative of the normal field n can be written in terms of the acceleration and the Weingarten map as follows

$$\nabla n = \varepsilon n^\flat \otimes a + \textbf{Wein}\,. \qquad (78)$$

Recalling (77), the purely covariant version of this is

$$\nabla n^\flat = -\varepsilon\left(K - n^\flat \otimes a^\flat\right)\,. \qquad (79)$$

From (48) and (79) we derive by standard manipulation, using vanishing torsion,

$$L_n h = -2\varepsilon K\,. \qquad (80)$$

In presence of torsion there would be an additional term $+2(i_n T)^\flat_s$, where the subscript s denotes symmetrization; in coordinates $[(i_n T)^\flat_s]_{\mu\nu} = n^\lambda T^\alpha_{\lambda(\mu} g_{\nu)\alpha}$.

5   Curvature tensors

We wish to calculate the (intrinsic) curvature tensor of ∇ and express it in terms of the curvature tensor of D, the extrinsic curvature K, and the spatial and normal derivatives of n and K. Before we do this, we wish to say a few words on the definition of the curvature measures in general.

All notions of curvature eventually reduce to that of curves. For a surface S embedded in $\mathbb{R}^3$ we have the notion of Gaussian curvature which comes about as follows: Consider a point $p \in S$ and a unit vector v at p tangent to S. Consider all smooth curves passing through p with unit tangent v. It is easy to see that the curvatures at p of all such curves is not bounded from above (due to the possibility to bend within the surface), but there will be a lower bound, $k(p,v)$, which just depends on the chosen point p and the tangent direction represented by v. Now consider $k(p,v)$ as function of v. As v varies over all tangent directions $k(p,v)$ will

assume a minimal and a maximal value, denoted by $k_{min}(p) = k(p,v_{min})$ and $k_{max}(p) = k(p,v_{max})$ respectively. These are called the principal curvatures of S at p and their reciprocals are called the principal radii. It is clear that the principal directions v_{min} and v_{max} just span the eigenspaces of the Weingarten map discussed above. In particular, v_{min} and v_{max} are orthogonal. The Gaussian curvature $K(p)$ of S at p is then defined to be the product of the principal curvatures:

$$K(p) = k_{min}(p) \cdot k_{max}(p)\,. \qquad (81)$$

This definition is extrinsic in the sense that essential use is made of the ambient $\mathbb{R}^3$ in which S is embedded. However, Gauss' *theorema egregium* states that this notion of curvature can also be defined intrinsically, in the sense that the value $K(p)$ can be obtained from geometric operations entirely carried out *within* the surface S. More precisely, it is a function of the first fundamental form (the metric) only, which encodes the intrinsic geometry of S, and does not involve the second fundamental form (the extrinsic curvature), which encodes how S is embedded into $\mathbb{R}^3$.

Let us briefly state Gauss' theorem in mathematical terms. Let

$$g = g_{ab}\, dx^a \otimes dx^b \qquad (82)$$

be the metric of the surface in some coordinates, and

$$\Gamma^c_{ab} = \tfrac{1}{2} g^{cd}\left(-\partial_d g_{ab} + \partial_a g_{bd} + \partial_b g_{da}\right), \qquad (83)$$

certain combinations of first derivatives of the metric coefficients, known under the name of *Christoffel symbols* . Note that Γ^c_{ab} has as many independent components as $\partial_a g_{bc}$ and that we can calculate the latter from the former via

$$\partial_c g_{ab} = g_{an}\Gamma^n_{bc} + g_{bn}\Gamma^n_{ac}\,. \qquad (84)$$

Next we form even more complicated combinations of first and second derivatives of the metric coefficients, namely

$$R^a_{\ b\,cd} = \partial_c \Gamma^a_{db} - \partial_d \Gamma^a_{cb} + \Gamma^a_{cn}\Gamma^n_{db} - \Gamma^a_{dn}\Gamma^n_{cb}\,, \quad (85)$$

15

which are now known as components of the *Riemann curvature tensor*. From them we form the totally covariant (all indices down) components:

$$R_{ab\,cd} = g_{an} R^n{}_{b\,cd}. \tag{86}$$

They are antisymmetric in the first and second index pair: $R_{ab\,cd} = -R_{ba\,cd} = -R_{ab\,dc}$, so that $R_{12\,12}$ is the only independent component. Gauss' theorem now states that at each point on S we have

$$K = \frac{R_{12\,12}}{g_{11}g_{22} - g_{12}^2}. \tag{87}$$

An important part of the theorem is to show that the right-hand side of (87) actually makes good geometric sense, i.e. that it is independent of the coordinate system that we use to express the coefficients. This is easy to check once one knows that R_{abcd} are the coefficients of a tensor with the symmetries just stated. In this way the curvature of a surface, which was primarily defined in terms of curvatures of certain curves on the surface, can be understood intrinsically. In what follows we will see that the various measures of intrinsic curvatures of n-dimensional manifolds can be reduced to that of 2-dimensional submanifolds, which will be called *sectional curvatures*.

Back to the general setting, we start from the notion of a covariant derivative ∇. Its associated curvature tensor is defined by

$$R(X,Y)Z = \left(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\right)Z. \tag{88}$$

For each point $p \in M$ it should be thought of as a map that assigns to each pair $X,Y \in T_pM$ of tangent vectors at p a linear map $R(X,Y) : T_pM \to T_pM$. This assignment is antisymmetric, i.e. $R(X,Y) = -R(Y,X)$. If $R(X,Y)$ is applied to Z the result is given by the right-hand side of (88). Despite first appearance, the right-hand side of (88) at a point $p \in M$ only depends on the values of X, Y, and Z at that point and hence defines a tensor field. This one again proves by showing the validity of $R(fX,Y)Z = R(X,fY)Z = R(X,Y)fZ = fR(X,Y)Z$ for all smooth real-valued functions f on M. In other words: All terms involving derivatives of f cancel.

From (88) and using (71) one may show that the Riemann tensor always obeys the first and second *Bianchi identities*:

$$\sum_{(XYZ)} R(X,Y)Z$$
$$= \sum_{(XYZ)} \left\{ (\nabla_X T)(Y,Z) - T\big(X,T(Y,Z)\big) \right\}, \tag{89a}$$

$$\sum_{(XYZ)} (\nabla_X R)(Y,Z)$$
$$= \sum_{(XYZ)} R\big(X, T(Y,Z)\big), \tag{89b}$$

where the sums are over the three cyclic permutations of X, Y, and Z. For zero torsion these identities read in component form:

$$\sum_{(\lambda\mu\nu)} R^\alpha{}_{\lambda\,\mu\nu} = 0, \tag{90a}$$

$$\sum_{(\lambda\mu\nu)} \nabla_\lambda R^\alpha{}_{\beta\,\mu\nu} = 0. \tag{90b}$$

The second traced on (α,μ) and contracted with $g^{\beta\nu}$ yields (-2) times (13).

The covariant Riemann tensor is defined by

$$\mathbf{Riem}(W,Z,X,Y) := g\big(W, R(X,Y)Z\big). \tag{91}$$

For general covariant derivatives its only symmetry is the antisymmetry in the last pair. But for special choices it acquires more. In standard GR we assume the covariant derivative to be *metric compatible* and *torsion free*:

$$\nabla g = 0, \tag{92}$$
$$T = 0. \tag{93}$$

In that case the Riemann tensor has the symmetries

$$\mathbf{Riem}(W,Z,X,Y) = -\mathbf{Riem}(W,Z,Y,X), \tag{94a}$$
$$\mathbf{Riem}(W,Z,X,Y) = -\mathbf{Riem}(Z,W,X,Y), \tag{94b}$$
$$\mathbf{Riem}(W,X,Y,Z) + \mathbf{Riem}(W,Y,Z,X) +$$
$$\mathbf{Riem}(W,Z,Y,X) = 0, \tag{94c}$$
$$\mathbf{Riem}(W,Z,X,Y) = \mathbf{Riem}(X,Y,W,Z). \tag{94d}$$

Equation (94a) is true by definition (88), (94b) is equivalent to metricity of ∇, and (94c) is the first Bianchi identity in case of zero torsion. The last symmetry (94d) is a consequence of the preceding three. Together (94a), (94b), and (94d) say that, at each point $p \in M$, **Riem** can be thought of as symmetric bilinear form on the antisymmetric tensor product $T_pM \wedge T_pM$. The latter has dimension $N = \frac{1}{2}n(n-1)$ if M has dimension n, and the space of symmetric bilinear forms has dimension $\frac{1}{2}N(N+1)$. From that number we have to subtract the number of independent conditions (94c), which is $\binom{n}{4}$ in dimensions $n \geq 4$ and zero otherwise. Indeed, it is easy to see that (94c) is identically satisfied as a consequence of (94a) and (94b) if any two vectors W, Z, X, Y coincide (proportionality is sufficient). Hence the number $\#$ of independent components of the curvature tensor is

$\#\mathbf{Riem} =$
$$\begin{cases} \frac{1}{2}N(N+1) - \binom{n}{4} = \frac{1}{12}n^2(n^2-1) & \text{for} \quad n \geq 4 \\ 6 & \text{for} \quad n = 3 \\ 1 & \text{for} \quad n = 2 \end{cases}$$
$$= \tfrac{1}{12}n^2(n^2-1) \quad \text{for all } n \geq 2 \,.$$
$$(95)$$

The Ricci and scalar curvatures are obtained by taking traces with respect to g: Let $\{e_1, \cdots, e_n\}$ be an orthonormal basis, $g(e_a, e_b) = \delta_{ab}\varepsilon_a$ (no summation) with $\varepsilon_a = \pm 1$, then

$$\mathbf{Ric}(X, Y) = \sum_{a=1}^{n} \varepsilon_a \, \mathbf{Riem}(e_a, X, e_a, Y) \qquad (96)$$

$$\mathbf{Scal} = \sum_{a=1}^{n} \varepsilon_a \, \mathbf{Ric}(e_a, e_a) \,. \qquad (97)$$

The Einstein tensor is

$$\mathbf{Ein} = \mathbf{Ric} - \tfrac{1}{2}\mathbf{Scal}\, g \,. \qquad (98)$$

The sectional curvature is defined by

$$\mathbf{Sec}(X, Y) = \frac{\mathbf{Riem}(X, Y, X, Y)}{g(X,X)g(Y,Y) - \big[g(X,Y)\big]^2} \,, \quad (99)$$

Here X, Y is a pair of linearly independent tangent vectors that span a 2-dimensional tangent subspace restricted to which g is non-degenerate. We will say that span$\{X, Y\}$ is non-degenerate. This is the necessary and sufficient condition for the denominator on the right-hand side to be non zero. The quantity $\mathbf{Sec}(X, Y)$ is called the *sectional curvature* of the manifold (M, g) at point p tangent to span$\{X, Y\}$. From the symmetries of **Riem** it is easy to see that the right-hand side of (99) does indeed only depend on the span of X, Y. That is, for any other pair X', Y' such that span$\{X', Y'\} =$ span$\{X, Y\}$, we have $\mathbf{Sec}(X', Y') = \mathbf{Sec}(X, Y)$. The geometric interpretation of $\mathbf{Sec}(X, Y)$ is as follows: Consider all geodesics of (M, g) that pass through the considered point $p \in M$ in a direction tangential to span$\{X, Y\}$. In a neighborhood of p they form an embedded 2-surface in M whose Gaussian curvature is just $\mathbf{Sec}(X, Y)$.

Now, **Riem** is determined by components of the form $\mathbf{Riem}(X, Y, X, Y)$, as follows from the fact that **Riem** is a *symmetric* bilinear form on $TM \wedge TM$. This remains true if we restrict to those X, Y whose span is non-degenerate, since they lie dense in $TM \wedge TM$ and $(X, Y) \mapsto \mathbf{Riem}(X, Y, X, Y)$ is continuous. This shows that the full information of the Riemann tensor can be reduced to certain Gaussian curvatures.

This also provides a simple geometric interpretation of the scalar and Einstein curvatures in terms of sectional curvatures. Let $\{X_1, \cdots, X_n\}$ be any set of pairwise orthogonal non-null vectors. The $\frac{1}{2}n(n-1)$ 2-planes span$\{X_a, X_b\}$ are non-degenerate and also pairwise orthogonal. It then follows from (97) and (99) that the scalar curvature is twice the sum of all sectional curvatures:

$$\mathbf{Scal} = 2 \sum_{\substack{a,b=1 \\ a<b}}^{n} \mathbf{Sec}(X_a, X_b) \,. \qquad (100)$$

The sum on the right-hand side of (100) is the same for any set of $\frac{1}{2}n(n-1)$ non-degenerate and pairwise orthogonal 2-planes. Hence the scalar curvature can be said to be twice the sum of mutually orthogonal sectional curvatures, or $n(n-1)$ times the mean sectional curvature. Similarly for the Ricci

and Einstein curvatures. The symmetry of the Ricci and Einstein tensors imply that they are fully determined by their components $\mathbf{Ric}(W, W)$ and $\mathbf{Ein}(W, W)$. Again this remains true if we restrict to the dense set of non-null W, i.e. $g(W, W) \neq 0$. Let now $\{X_1, \cdots, X_{n-1}\}$ be any set of mutually orthogonal vectors (again they need not be normalized) in the orthogonal complement of W. As before the $\frac{1}{2}(n-1)(n-2)$ planes $\mathrm{span}\{X_a, X_b\}$ are non degenerate and pairwise orthogonal. From (96), (98), and (99) it follows that

$$\mathbf{Ric}(W, W) = g(W, W) \sum_{a=1}^{n-1} \mathbf{Sec}(W, X_a) \quad (101)$$

and

$$\mathbf{Ein}(W, W) = -g(W, W) \sum_{\substack{a,b=1 \\ a<b}}^{n-1} \mathbf{Sec}(X_a, X_b). \quad (102)$$

Again the right-hand sides will be the same for any set $\{X_1, \cdots, X_{n-1}\}$ of $n-1$ mutually orthogonal vectors in the orthogonal complement of W. Note that $\mathbf{Ric}(W, W)$ involves all sectional curvatures involving W whereas $\mathbf{Ein}(W, W)$ involves all sectional curvatures orthogonal to W. For normalized W, where $g(W, W) = \sigma = \pm 1$, we can say that $-\sigma G(W, W)$ is the sum of sectional curvatures orthogonal to W, or $\frac{1}{2}(n-1)(n-2)$ times their mean. Note that for timelike W we have $\sigma = -1$ and $G(W, W)$ is just the sum of spatial sectional curvatures.

Finally we mention the Weyl curvature tensor, which contains that part of the information in the curvature tensor not captured by the Ricci (or Einstein-) tensor. To state its form in a compact form, we introduce the *Kulkarni-Nomizu product*, denoted by an encircled wedge, $\owedge$, which is a bilinear symmetric product on the space of covariant symmetric rank-two tensors with values in the covariant rank-four tensors that have the symmetries (94) of the Riemann tensor. Let k and ℓ be two symmetric covariant second-rank tensors, then

their Kulkarni-Nomizu product is defined by

$$\begin{aligned} k \owedge \ell(X_1, X_2, X_3, X_4) := & k(X_1, X_3)\,\ell(X_2, X_4) \\ & + k(X_2, X_4)\,\ell(X_1, X_3) \\ & - k(X_1, X_4)\,\ell(X_2, X_3) \\ & - k(X_2, X_3)\,\ell(X_1, X_4), \end{aligned}$$
$$(103)$$

or in components

$$(k \owedge \ell)_{abcd} = k_{ac}\ell_{bd} + k_{bd}\ell_{ac} - k_{ad}\ell_{bc} - k_{bc}\ell_{ad}. \quad (104)$$

The Weyl tensor, $\mathbf{Weyl}$, is of the same type as $\mathbf{Riem}$ but in addition totally trace-free. It is obtained from $\mathbf{Riem}$ by a projection map, P_W, given by

$$\begin{aligned} \mathbf{Weyl} := & P_W(\mathbf{Riem}) \\ := & \mathbf{Riem} - \tfrac{1}{n-2}\Big(\mathbf{Ric} - \tfrac{1}{2(n-1)}\mathbf{Scal}\,g\Big) \owedge g. \end{aligned}$$
$$(105)$$

P_W is a linear map from the space of rank-four tensors with Riemann symmetries to itself. It is easy to check that its image is given by the totally trace-free such tensors and that the kernel consists of all tensors of the form $g \owedge K$, where K is a symmetric rank-two tensor. The latter clearly implies $P_W \circ P_W = P_W$. The dimension of the image corresponds to the number of independent components of the Weyl tensor, which is given by (95) minus the dimension $\frac{1}{2}n(n+1)$ of the kernel. This gives for $n \geq 3$

$$\#\mathbf{Weyl} = \tfrac{1}{12}n(n+1)\big[n(n-1) - 6\big] \quad (106)$$

and zero for $n = 2$. Note that in $n = 3$ dimensions the Weyl tensor also always vanishes, so that (105) can be used to express the Riemann tensor in terms of the Ricci and scalar curvature

$$\mathbf{Riem} = \big(\mathbf{Ric} - \tfrac{1}{4}\mathbf{Scal}\,g\big) \owedge g \quad \text{(for } n = 3\text{)}. \quad (107)$$

A metric manifold (M, g) is said to be of constant curvature if

$$\mathbf{Riem} = k\,g \owedge g \quad (108)$$

18

for some function k. Then $\mathbf{Ric} = 2k(n-1)g$ and $\mathbf{Ein} = -k(n-1)(n-2)g$. We recall that manifolds (M, g) for which the Einstein tensor (equivalently, the Ricci tensor) is pointwise proportional to the metric are called *Einstein spaces*. The twice contracted second Bianchi identity (13) shows that k must be a constant unless $n = 2$. For $n = 3$ equation (107) shows that Einstein spaces are of constant curvature.

5.1 Comparing curvature tensors

Sometimes one wants to compare two different curvature tensors belonging to two different covariant derivatives $\hat{\nabla}$ and ∇. In what follows, all quantities referring to $\hat{\nabla}$ carry a hat. Recall that a covariant derivative can be considered as a map $\nabla : \Gamma T_0^1 M \times \Gamma T_0^1 M \to \Gamma T_0^1 M$, $(X, Y) \mapsto \nabla_X Y$, which is $C^\infty(M)$-linear in the first and a derivation in the second argument. That is, for $f \in C^\infty(M)$ have $\nabla_{fX+Y} Z = f \nabla_X Z + \nabla_Y Z$ and $\nabla_X(fY + Z) = X(f)Y + f \nabla_X Y + \nabla_X Z$. This implies that the difference of two covariant derivatives is $C^\infty(M)$- linear also in the second argument and hence a tensor field:

$$\hat{\nabla} - \nabla =: \Delta \in \Gamma T_2^1 M \, . \tag{109}$$

Replacing $\hat{\nabla}$ with $\nabla + \Delta$ in the definition of the curvature tensor for $\hat{\nabla}$ according to (88) directly leads to

$$\begin{aligned}
\hat{R}(X, Y)Z = {}& R(X, Y)Z \\
& + (\nabla_X \Delta)(Y, Z) - (\nabla_Y \Delta)(X, Z) \\
& + \Delta\big(X, \Delta(Y, Z)\big) - \Delta\big(Y, \Delta(X, Z)\big) \\
& + \Delta\big(T(X, Y), Z\big) \, .
\end{aligned} \tag{110}$$

Note that so far no assumptions have been made concerning torsion or metricity of $\hat{\nabla}$ and ∇. This formula is generally valid. In the special case where $\hat{\nabla}$ and ∇ are the Levi-Civita covariant derivatives with respect to two metrics $\hat{g}$ and g, we set

$$h := \hat{g} - g \, , \tag{111}$$

which is a symmetric covariant tensor field. Note that here, and for the rest of this subsection, h has

a different meaning from that given to it in (48). We recall that the Levi-Civita covariant derivative is uniquely determined by the metric. For ∇ this reads

$$\begin{aligned}
2\, g(\nabla_X Y, Z) \\
= X\big(g(Y, Z)\big) + Y\big(g(Z, X)\big) - Z\big(g(X, Y)\big) \\
- g\big(X, [Y, Z])\big]\big) + g\big(Y, [Z, X])\big]\big) + g\big(Z, [X, Y])\big]\big).
\end{aligned} \tag{112}$$

Subtracting (112) from the corresponding formula with ∇ and g replaced by $\hat{\nabla}$ and $\hat{g}$ yields, using $T = 0$,

$$\begin{aligned}
2\, \hat{g}\big(\Delta(X, Y), Z\big) = \\
- (\nabla_Z h)(X, Y) + (\nabla_X h)(Y, Z) + (\nabla_Y h)(Z, X).
\end{aligned} \tag{113}$$

This formula expresses Δ as functional of g and $\hat{g}$. There are various equivalent forms of it. We have chosen a representation that somehow minimizes the appearance of $\hat{g}$. Note that g enters in h as well as ∇, whereas $\hat{g}$ enters in h and via the scalar product on the left-hand side. The latter obstructs expressing Δ as functional of g and h alone. In components (113) reads

$$\Delta_{bc}^a = \tfrac{1}{2} \hat{g}^{an} \big(-\nabla_n h_{bc} + \nabla_b h_{cn} + \nabla_c h_{nb} \big) . \tag{114}$$

Note that one could replace the components of h with those of $\hat{g} = g + h$ in the bracket on the right-hand side, since the covariant derivatives of g vanish.

Now suppose we consider h and its first and second derivatives to be small and we wanted to know the difference in the covariant derivatives and curvature only to leading (linear) order in h. To that order we may replace $\hat{g}$ with g on the left-hand side of (113) and the right-hand side of (114). Moreover we may neglect the Δ-squared terms in (110) and obtain, writing δR for the first order contribution to $\hat{R} - R$,

$$\delta R^a{}_{bcd} = \nabla_c \Delta_{db}^a - \nabla_d \Delta_{cb}^a . \tag{115}$$

From this the first-order variation of the Ricci ten-

sor follows, writing $h_{ab} =: \delta g_{ab}$,

$$\begin{aligned} \delta R_{ab} &= \nabla_n \Delta_{ab}^n - \nabla_b \Delta_{na}^n \\ &= \tfrac{1}{2}\big(-\Delta_g\,\delta g_{ab} - \nabla_a\nabla_b\,\delta g \\ &\quad + \nabla^n\nabla_a\,\delta g_{nb} + \nabla^n\nabla_b\,\delta g_{na}\big)\,, \end{aligned} \tag{116}$$

where $\Delta_g := g^{ab}\nabla_a\nabla_b$ and $\delta g = g^{ab}\,\delta g_{ab}$. Finally, the variation of the scalar curvature is (note $\delta g^{ab} = -g^{ac}g^{bd}\delta g_{cd} = -h^{ab}$)

$$\delta R = R_{ab}\,\delta g^{ab} + \nabla_a U^a\,, \tag{117a}$$

where

$$\begin{aligned} U^a &= g^{nm}\Delta_{nm}^a - g^{an}\Delta_{mn}^m \\ &= G^{abcd}\nabla_b\,\delta g_{cd}\,. \end{aligned} \tag{117b}$$

Here we made use of the *De Witt metric*, which defines a symmetric non-degenerate bilinear form on the space of symmetric covariant rank-two tensors and which in components reads:

$$G^{abcd} = \tfrac{1}{2}\big(g^{ac}g^{bd} + g^{ad}g^{bc} - 2g^{ab}g^{cd}\big)\,. \tag{118}$$

We will later have to say more about it.

We also wish to state a useful formula that compares the curvature tensors for conformally related metrics, i.e.

$$\hat{g} = e^{2\phi}\,g\,, \tag{119}$$

where $\phi : M \to \mathbb{R}$ is smooth. Then

$$\mathbf{Riem}_{\hat{g}} = e^{2\phi}\Big[\mathbf{Riem}_g + g \oslash K\Big]\,, \tag{120a}$$

with

$$K := -\nabla^2\phi + d\phi \otimes d\phi - \tfrac{1}{2}g^{-1}(d\phi, d\phi)\,g\,. \tag{120b}$$

(This can be proven by straightforward calculations using either (88) and (112), or Cartan's structure equations, or, most conveniently, normal coordinates.) From (120a) and the fact that the kernel of the map P_W in (105) is given by tensors of the form $g \oslash K$ it follows immediately that

$$\mathbf{Weyl}_{\hat{g}} = e^{2\phi}\mathbf{Weyl}_g\,. \tag{121}$$

This is equivalently expressed by the conformal invariance of the contravariant version of the Weyl tensor, which is related to the covariant form, **Weyl**, in the same way (91) as the curvature tensor R is related to **Riem** (i.e., by raising the first index of the latter).

From (120) we also deduce the transformation properties of the Ricci tensor:

$$\begin{aligned} \mathbf{Ric}_{\hat{g}} &= \mathbf{Ric}_g \\ &\quad - \big(\Delta_g\phi + (n-2)g^{-1}(d\phi, d\phi)\big)\,g \\ &\quad - (n-2)(\nabla\nabla\phi - d\phi \otimes d\phi)\,. \end{aligned} \tag{122}$$

where, as above, Δ_g denotes again the Laplacian/d'Alembertian for g. Finally, for the scalar curvature we get

$$\begin{aligned} \mathbf{Scal}_{\hat{g}} = e^{-2\phi}\bigg(&\mathbf{Scal}_g \\ &- 2(n-1)\Delta_g\phi \\ &- (n-1)(n-2)g^{-1}(d\phi, d\phi)\bigg)\,. \end{aligned} \tag{123}$$

This law has a linear dependence on the second and a quadratic dependence on first derivatives of ϕ. If the conformal factor is written as an appropriate power of some positive function $\Omega : M \to \mathbb{R}_+$ we can eliminate all dependence on first and just retain the second derivatives. In $n > 2$ dimensions it is easy to check that the rule is this:

$$e^{2\phi} = \Omega^{\frac{4}{n-2}}\,, \tag{124}$$

then (123) becomes

$$\mathbf{Scal}_{\hat{g}} = -\frac{4(n-1)}{n-2}\Omega^{-\frac{n+2}{n-2}}\mathcal{D}_g\Omega\,, \tag{125a}$$

where

$$\mathcal{D}_g = \Delta_g - \frac{n-2}{4(n-1)}\mathbf{Scal}_g\,. \tag{125b}$$

$\mathcal{D}_g$ is a linear differential operator which is elliptic for Riemannian and hyperbolic for Lorentzian metrics g. If we set $\Omega = \Omega_1\Omega_2$ and apply (125) twice, one time to the pair $(\hat{g}, g)$, the other time to $(\hat{g}, \Omega_2 g)$, we obtain by direct comparison (and

renaming Ω_2 to Ω thereafter) the conformal transformation property for the operator $\mathcal{D}_g$:

$$\mathcal{D}_{\Omega^{\frac{4}{n-2}}g} = M\left(\Omega^{-\frac{n+2}{n-2}}\right) \circ \mathcal{D}_g \circ M(\Omega), \quad (126)$$

where $M(\Omega)$ is the linear operator of multiplication with Ω. This is the reason why $\mathcal{D}_g$ is called the *conformally covariant Laplacian* (for Riemannian g) or the *conformally covariant wave operator* (for Lorentzian g). As we will see, it has useful applications to the initial-data problem in GR.

5.2 Curvature decomposition

Using (67) we can decompose the various curvature tensors. First we let X, Y, Z be horizontal vector fields. We use (67) in (88) and get the general formula (i.e. not yet making use of the fact that ∇ and D are metric and torsion free)

$$\begin{aligned} R(X,Y)Z = R^D&(X,Y)Z \\ &+ (\nabla_X n)\,K(Y,Z) - (\nabla_Y n)\,K(X,Z) \\ &+ n\big[(D_X K)(Y,Z) - (D_Y K)(X,Z)\big] \\ &+ n\,K\big(T^D(X,Y),Z\big), \end{aligned}$$
$$(127)$$

where

$$R^D(X,Y,)Z := \big(D_X D_Y - D_Y D_X - D_{[X,Y]}\big)Z \tag{128}$$

is the horizontal curvature tensor associated to the Levi-Civita covariant derivative D of h. This formula is general in the sense that it is valid for any covariant derivative. No assumptions have been made so far concerning metricity or torsion, and this is why the torsion T^D of D (defined in (70)) makes an explicit appearance. From now on we shall restrict to vanishing torsion. We observe that the first two lines on the right-hand side of (127) are horizontal whereas the last two lines are proportional to n. Decomposition into horizontal and normal components, respectively, leads to (where $T^D = 0$ and X, Y, Z, and W are horizontal),

$$\begin{aligned} \mathbf{Riem}(W,Z,X,Y) &= \mathbf{Riem}^D(W,Z,X,Y) \\ &- \varepsilon\big[K(W,X)K(Z,Y) - K(W,Y)K(Z,X)\big]. \end{aligned}$$
$$(129)$$

Here we used $h(W, \nabla_X n) = -\epsilon K(W,X)$ from (75), and

$$\begin{aligned} \mathbf{Riem}&(n,Z,X,Y) \\ &= \varepsilon\big[(D_X K)(Y,Z) - (D_Y K)(X,Z)\big]. \end{aligned}$$
$$(130)$$

Here and in the sequel we return to the meaning of h given by (48). In differential geometry (129) is referred to as *Gauss equation* and (130) as *Codazzi-Mainardi equation*.

The remaining curvature components are those involving two entries in n direction. Using (79) we obtain via standard manipulations (now using metricity and vanishing torsion)

$$\begin{aligned} \mathbf{Riem}&(X,n,Y,n) \\ &= i_X\big(\nabla_Y \nabla_n - \nabla_n \nabla_Y - \nabla_{[Y,n]}\big)n^\flat \\ &= i_X i_Y\big(\varepsilon L_n K + K \circ K + Da^\flat - \varepsilon a^\flat \otimes a^\flat\big). \end{aligned}$$
$$(131)$$

Here $K \circ K(X,Y) := h^{-1}(i_X K, i_Y K) = i_X K\big((i_Y K)^\sharp\big)$ and we used the following relation between covariant and Lie derivative (which will have additional terms in case of non-vanishing torsion):

$$\nabla_n K = L_n K + 2\varepsilon\,K \circ K\,. \tag{132}$$

Note also that the left-hand side of (131) is symmetric as consequence of (94d). On the right-hand side only $Da^\flat$ is not immediately seen to be symmetric, but that follows from (52b). Unlike (129) and (130), equation (131) does not seem to have a standard name in differential geometry.

Equations (127), (129), and (130) express all components of the spacetime curvature in terms of horizontal quantities and their Lie derivatives L_n in normal direction. According to (55) the latter can be replaced by a combination of Lie derivatives along the time vector-field $\partial/\partial t$ and the shift β. From (53b) we infer that $L_{\alpha n} = \alpha L_n$ on horizontal covariant tensor fields, therefore we may replace

$$L_n \to \alpha^{-1}\big(L_{\frac{\partial}{c\partial t}} - L_\beta\big) \to \alpha^{-1}\big(L^{\|}_{\frac{\partial}{c\partial t}} - L^{\|}_\beta\big) \tag{133}$$

on horizontal covariant tensor fields. Here we set $L^{\|} = P^{\|} \circ L$, i.e. Lie derivative (as operation in the

ambient spacetime) followed by horizontal projection. Moreover, using (50), one easily sees that the acceleration 1-form $a^\flat$ can be expressed in terms of the spatial derivative of the lapse function:

$$a^\flat = -\varepsilon\alpha^{-1}D\alpha\,. \qquad (134)$$

Hence the combination of accelerations appearing in (130) may be written as

$$Da^\flat - \varepsilon a^\flat \otimes a^\flat = -\varepsilon\alpha^{-1}D^2\alpha\,. \qquad (135)$$

Note that $D^2\alpha := DD\alpha$ is just the horizontal covariant Hessian of α with respect to h.

6 Decomposing Einstein's equations

The curvature decomposition of the previous section can now be used to decompose Einstein's equations. For this we decompose the Einstein tensor **Ein** into the normal-normal, normal-tangential, and tangential-tangential parts. Let $\{e_0, e_1, e_2, e_3\}$ be an orthonormal frame with $e_0 = n$, i.e. adapted to the foliation as in Section 4.1. Then (102) together with (129) immediately lead to

$$2\,\mathbf{Ein}(e_0, e_0) = -\left[K_{ab}K^{ab} - (K_a^a)^2\right] - \varepsilon\mathbf{Scal}^D\,, \qquad (136)$$

where $\mathbf{Scal}^D$ is the scalar curvature of D, i.e. of the spacelike leaves in the metric h. Similarly we obtain from (130),

$$\mathbf{Ein}(e_0, e_a) = \mathbf{Ric}(e_0, e_a) = -\varepsilon\left[D^bK_{ab} - D_aK_b^b\right]\,. \qquad (137)$$

The normal-normal component of the Ricci tensor cannot likewise be expressed simply in terms of horizontal quantities, the geometric reason being that, unlike the Einstein tensor, it involves non-horizontal sectional curvatures (compare (101) and (102)). A useful expression follows from taking the trace of (131), considered as symmetric bilinear form in X and Y. The result is :

$$\mathbf{Ric}(e_0, e_0) = -K_{ab}K^{ab} + (K_c^c)^2 + \varepsilon\nabla\cdot V\,, \qquad (138)$$

where $\nabla\cdot$ denotes the divergence with respect to ∇ and V is a vector field on M whose normal component is the trace of the extrinsic curvature and whose horizontal component is ε times the acceleration on n:

$$V = nK_c^c + \varepsilon a\,. \qquad (139)$$

For the horizontal-horizontal components of Einstein's equation it turns out to be simpler to use their alternative form (6b) with the Ricci tensor on the left hand side. For that we need the horizontal components of the Ricci tensor, which we easily get from (129) and (131):

$$\begin{aligned}
\mathbf{Ric}(e_a, e_b) = \,&\mathbf{Ric}^D(e_a, e_b) \\
&+ L_nK_{ab} + 2\varepsilon K_{ac}K_b^c - \varepsilon K_{ab}K_c^c \\
&+ \varepsilon D_aa_b - a_aa_b\,.
\end{aligned} \qquad (140)$$

For later applications we also note the expression for the scalar curvature. It follows, e.g., from adding the horizontal trace of (140) to ε times (138). This leads to

$$\mathbf{Scal} = \mathbf{Scal}^D - \varepsilon\left[K_{ab}K^{ab} - (K_a^a)^2\right] + 2\nabla\cdot V\,. \qquad (141)$$

Here we made use of the relation between the ∇ and D derivative for the acceleration 1-form:

$$\nabla a^\flat = Da^\flat + \varepsilon\,n^\flat \otimes \nabla_n a^\flat + i_a K \otimes n^\flat\,, \qquad (142)$$

whose trace gives the following relation between the ∇ and D divergences of a:

$$\nabla\cdot a = D\cdot a - \varepsilon h(a, a)\,. \qquad (143)$$

Another possibility would have been to use (136) and (138) in $\mathbf{Scal} = -2\varepsilon(\mathbf{Ein}(e_0, e_0) - \mathbf{Ric}(e_0, e_0))$.

Using (136) and (137), and also using the De Witt metric (118) for notational ease, we can immediately write down the normal-normal and normal-tangential components of Einstein's equations (3):

$$G^{abcd}K_{ab}K_{cd} + \varepsilon\mathbf{Scal}^D = -2\kappa\,\mathbf{T}(n, n)\,, \quad (144\text{a})$$

$$G^{abcd}D_bK_{cd} = -\varepsilon\kappa h^{ab}\mathbf{T}(n, e_b)\,. \qquad (144\text{b})$$

From (77) and (118) we notice that the bilinear form on the left-hand side of (144a) can be written as

$$G(K, K) := G^{abcd} K_{ab} K_{cd}$$
$$= \mathrm{Tr}(\mathbf{Wein} \circ \mathbf{Wein}) - \left(\mathrm{Tr}(\mathbf{Wein})\right)^2.$$
$$(145)$$

Here the trace is natural (needs no metric for its definition) since **Wein** is an endomorphism. In a local frame in which **Wein** is diagonal with entries $\vec{k} := (k_1, k_2, k_3)$ we have

$$G(K, K) := (\delta^{ab} - 3n^a n^b) k_a k_b, \qquad (146)$$

where n^a are the components of the normalized vector $(1, 1, 1)/\sqrt{3}$ in eigenvalue-space, which we identify with $\mathbb{R}^3$ endowed with the standard Euclidean inner product. Hence, denoting by θ the angle between $\vec{n}$ and $\vec{k}$, we have

$$G(K, K) = \begin{cases} 0 & \text{if } |\cos\theta| = \sqrt{1/3} \\ > 0 & \text{if } |\cos\theta| < \sqrt{1/3} \\ < 0 & \text{if } |\cos\theta| > \sqrt{1/3}. \end{cases} \quad (147)$$

Note that $|\cos\theta| = \sqrt{1/3}$ describes a double cone around the symmetry axis generated by $\vec{n}$ and vertex at the origin, whose opening angle just is right so as to contain all three axes of $\mathbb{R}^3$. For eigenvalue-vectors inside this cone the bilinear form is negative, outside this cone positive. Positive $G(K, K)$ require sufficiently anisotropic Weingarten maps, or, in other words, sufficiently large deviations from being umbilical points.

The horizontal-horizontal component of Einstein's equations in the form (5) immediately follows from (140). In the ensuing formula we use (133) to explicitly solve for the horizontal Lie derivative of K with respect to $\partial/c\partial t$ and also (135) to simplify the last two terms in (140). This results in

$$\dot{K}_{ab} := \left(L^{\|}_{\frac{\partial}{c\partial t}} K\right)_{ab}$$
$$= \left(L^{\|}_{\beta} K\right)_{ab} + D_a D_b \alpha$$
$$+ \alpha\left[-2\varepsilon K_{ac} K_b^c + \varepsilon K_{ab} K_c^c - \mathbf{Ric}^D(e_a, e_b)\right]$$
$$- \alpha\varepsilon \frac{\kappa}{n-2} h_{ab} \mathbf{T}(n, n)$$
$$+ \alpha\kappa\left(\mathbf{T} - \frac{1}{n-2} \mathrm{Tr}_h(\mathbf{T}) h\right)(e_a, e_b).$$
$$(148)$$

Note that in the last term the trace of $\mathbf{T}$ is taken with respect to h and not g. The relation is $\mathrm{Tr}_h(\mathbf{T}) = \mathrm{Tr}_g(\mathbf{T}) - \varepsilon\mathbf{T}(n, n)$.

The only remaining equation that needs to be added here is that which relates the time derivative of h with K. This we get from (80) and (133):

$$\dot{h}_{ab} := \left(L^{\|}_{\frac{\partial}{c\partial t}} h\right)_{ab} = \left(L^{\|}_{\beta} h\right)_{ab} - 2\alpha\epsilon K_{ab}. \quad (149)$$

Equations (149) and (148) are six first-order in time evolution equations for the pair (h, K). This pair cannot be freely specified but has to obey the four equations (144a) and (144b) which do not contain any time derivatives of h or K. Equations (144a) and (144b) are therefore referred to as *constraints*, more specifically (144a) as *scalar constraint* (also *Hamiltonian constraint*) and (144b) as *vector constraint* (also *diffeomorphism constraint*).

We derived these equations from the 3+1 split of a spacetime that we considered to be given. Despite having expressed all equations in terms of horizontal quantities, there is still a relic of the ambient space in our equations, namely the Lie derivative with respect to $\partial/\partial ct$. We now erase this last relic by interpreting this Lie derivative as ordinary partial derivative of some t-dependent tensor field on a genuine 3-dimensional manifold Σ, which is not thought of as being embedded into a spacetime. The horizontal projection $L^{\|}_{\beta}$ of the spacetime Lie derivative that appears on the right-hand sides of the evolution equations above then translates to the ordinary intrinsic Lie derivative on Σ with respect to β. This is how from now on we shall read the above equations. Spacetime does not yet

exist. Rather, it has to be constructed from the evolution of the fields according to the equations, usually complemented by the equations that govern the evolution of the matter fields. In these evolution equations α and β are freely specifiable functions, the choice of which is subject to mathematical/computational convenience. Once α and β are specified and h as a function of parameter-time has been determined, we can form the expression (64) for the spacetime metric and know that, by construction, it will satisfy Einstein's equations.

To sum up, the initial-value problem consists in the following steps:

1. Choose a 3-manifold Σ.

2. Choose a time-parameter dependent lapse function α and a time-parameter dependent shift vector-field β.

3. Find a Riemannian metric $h \in \Gamma T_2^0 \Sigma$ and a symmetric covariant rank-2 tensor field $K \in \Gamma T_2^0 \Sigma$ that satisfy equations (144a) and (144b) either in vacuum ($\mathbf{T} = \mathbf{T}_\Lambda$; cf. (4)), or after specifying some matter model.

4. Evolve these data via (149) and (148), possibly complemented by the evolution equations for the matter variables.

5. Construct from the solution the spacetime metric g via (64).

For this to be consistent we need to check that the evolution according to (149) and (148) will preserve the constraints (144a) and (144b). At this stage this could be checked directly, at least in the vacuum case. The easiest way to do this is to use the equivalence of these equations with Einstein's equations and then employ the twice contracted 2nd Bianchi identity (13). It follows that $\nabla_\mu E^{\mu\nu} \equiv 0$, where $E^{\mu\nu} = G^{\mu\nu} + \lambda g^{\mu\nu}$. The four constraints (144a) and (144b) are equivalent to $E^{00} = 0$ and $E^{0m} = 0$, and the six second-order equations $E^{mn} = 0$ to the twelve first-order evolution equations (149) and (148). In coordinates the identity $\nabla_\mu E^{\mu\nu} \equiv 0$ reads

$$\partial_0 E^{0\nu} = -\partial_m E^{m\nu} - \Gamma_{\mu\lambda}^\mu E^{\lambda\nu} - \Gamma_{\mu\lambda}^\nu E^{\mu\lambda}, \quad (150)$$

which shows immediately that the time derivatives of the constraint functions are zero if the constraints vanished initially. This suffices for analytic data, but in the general case one has to do more work. Fortunately the equations for the evolution of the constraint functions can be put into an equivalent form which is manifestly symmetric hyperbolic [62]. That suffices to conclude the preservation of the constraints in general. In fact, symmetric hyperbolicity implies more than that. It ensures the well-posedness of the initial-value problem for the constraints, which not only says that they stay zero if they are zero initially, but also that they stay small if they are small initially. This is of paramount importance in numerical evolution schemes, in which small initial violations of the constraints must be allowed for and hence the consequences of these violations need to be controlled. For a recent and mathematically more thorough discussion of the Cauchy problem we refer to James Isenberg's survey [86].

Finally we wish to substantiate our earlier claim that any Σ can carry some initial data. Let us show this for closed Σ. To this end we choose a matter model such that the right-hand side of (144b) vanishes. Note that this still allows for arbitrary cosmological constants since $\mathbf{T}_\Lambda(n, e_a) \propto g(n, e_a) = 0$. Next we restrict to those pairs (h, K) were $K = \lambda h$ for some constant λ. Geometrically this means that, in the spacetime to be developed, the Cauchy surface will be totally umbilical (isotropic Weingarten map). Due to this proportionality and the previous assumption the vector constraint (144b) will be satisfied. In the scalar constraint we have $G(K, K) = G(\lambda h, \lambda h) = -6\lambda^2$ so that it will be satisfied provided that

$$-\varepsilon \mathbf{Scal}^D = 2\kappa \mathbf{T}(n, n) - 6\lambda^2. \quad (151)$$

For the following argument the Lorentzian signature, $\varepsilon = -1$, will matter. For physical reasons we assume the weak energy condition so that $\kappa \mathbf{T}(n, n) \geq 0$, which makes a positive contribution to the right-hand side of (151). However, if we choose the modulus of λ sufficiently large we can make the right-hand side negative somewhere (or

everywhere, since Σ is compact). Now, in dimensions 3 or higher the following *theorem of Kazdan & Warner* holds ([91], Theorem 1.1): Any smooth function on a compact manifold which is negative somewhere is the scalar curvature for some smooth Riemannian metric. Hence a smooth h exists which solves (151) for any given $\mathbf{T}(n,n) \geq 0$, provided we choose $\lambda^2 > |\lambda|$ sufficiently large. If Σ is not closed a corresponding theorem may also be shown [118].

The above argument crucially depends on the signs. There is no corresponding statement for positive scalar curvatures. In fact, there is a strong topological obstruction against Riemannian metrics of strictly positive scalar curvature. It follows from the *theorem of Gromov & Lawson* ([78], Theorem 8.1) that a 3-dimensional closed orientable Σ allows for Riemannian metrics with positive scalar curvature iff its prime decomposition consists of prime-manifolds with finite fundamental group or "handles" $S^1 \times S^2$. All manifolds whose prime list contains at least one so-called $K(\pi,1)$-factor (a 3-manifold whose only non-trivial homotopy group is the first) are excluded. See, e.g., [67] for more explanation of these notions. We conclude that the given argument crucially depends on $\varepsilon = -1$.

6.1   A note on slicing conditions

The freedom in choosing the lapse and shift functions can be of much importance, theoretically and in numerical evolution schemes. This is particularly true for the lapse function α, which determines the amount of proper length by which the Cauchy slice advances in normal direction per unit parameter interval. If a singularity is to form in spacetime due to the collapse of matter within a bounded spatial region, it would clearly be advantageous to not let the slices run into the singularity before the outer parts of it have had any chance to develop a sufficiently large portion of spacetime that one might be interested in, e.g. for the study of gravitational waves produced in the past. This means that one would like to slow down α in regions which are likely to develop a singularity and speed up α in those regions where it seems affordable. Take as an example the "equal-speed" gauge $\alpha = 1$

and $\beta = 0$, so that $g = -c^2\,dt^2 + h$. This means that $n = \partial/\partial ct$ is geodesic. Taking such a gauge from the $t = 0$ slice in the Schwarzschild/Kruskal spacetime would let the slices run into the singularity after a proper-time of $t = \pi GM/c^3$, where M is the mass of the black hole. In that short period of time the slices had no chance to explore a significant portion of spacetime outside the black hole.

A gauge condition that one may anticipate to have singularity-avoiding character is that where α is chosen such that the divergence of the normal field n is zero. This condition just means that the locally co-moving infinitesimal volume elements do not change volume, for $L_n d\mu = (\nabla \cdot n)\,d\mu$, where $d\mu = \det\{h_{ab}\}d^3x$ is the volume element of Σ. From (79) we see that n has zero divergence iff K has zero trace, i.e. the slices are of zero mean-curvature. The condition on α for this to be preserved under evolution follows from

$$0 = L_n(h^{ab}K_{ab}) = -K^{ab}L_n h_{ab} + h^{ab}L_n K_{ab}\,. \tag{152}$$

Here we use (80) to eliminate $L_n h_{ab}$ in the first term and (131) to eliminate $L_n K_{ab}$ in the second term, also making use of (135). This leads to the following equivalent of (152):

$$\Delta_h \alpha + \varepsilon\big(\mathbf{Ric}(n,n) + K^{ab}K_{ab}\big)\alpha = 0\,. \tag{153}$$

This is a linear elliptic equation for α. The case of interest to us in GR is $\varepsilon = -1$. In the closed case we immediately deduce by standard arguments that $\alpha = 0$ is the only solution, provided the strong energy-condition holds (which implies $\mathbf{Ric}(n,n) \geq 0$). In the open case, where we might impose $\alpha \to 1$ as asymptotic condition, we deduce existence and uniqueness again under the assumption of the strong energy condition. Hence we may indeed impose the condition $h^{ab}K_{ab} = 0$, or $\mathrm{Tr}(\mathbf{Wein}) = 0$, for non-closed Σ. It is called the *maximal slicing condition* or *York gauge* [119].

Whereas this gauge condition has indeed the desired singularity-avoiding character it is also not easy to implement due to the fact that at each new stage of the evolution one has to solve the elliptic equation (153). For numerical studies it is easier

to implement evolution equations for α. Such an equation is, e.g., obtained by asking the time function (36) to be harmonic, in the sense that

$$0 = \Box_g t := g^{\mu\nu}\nabla_\mu\nabla_\nu t$$
$$= |\det\{g_{\alpha\beta}\}|^{-\frac{1}{2}}\partial_\mu\Big(|\det\{g_{\alpha\beta}\}|^{\frac{1}{2}}g^{\mu\nu}\partial_\nu\Big)t\,. \quad (154)$$

This is clearly just equivalent to

$$\partial_\mu\Big(|\det\{g_{\mu\nu}\}|^{\frac{1}{2}}g^{\mu 0}\Big) = 0\,, \quad (155)$$

which can be rewritten using (65) and (66) to give

$$\dot{\alpha} := \frac{\partial\alpha}{c\partial t} = L_\beta\alpha - \varepsilon K^a_a\alpha^2$$
$$= L_\beta\alpha + \mathrm{Tr}(\mathbf{Wein})\,\alpha^2\,. \quad (156)$$

This is called the *harmonic slicing condition*. Note that we can still choose $\beta = 0$ and try to determine α as function of the trace of **Wein**. There also exist generalizations to this condition where α^2 on the right-hand side is replaced with other functions $f(\alpha)$.

6.2   A note on the De Witt metric

At each point p on Σ the De Witt metric (118) can be regarded as a symmetric bilinear form on the space of positive-definite inner products h of $T_p\Sigma$. The latter is an open convex cone in $T_P^*\Sigma \otimes T_P^*\Sigma$. We wish to explore its properties a little further.

A frame in $T_p\Sigma$ induces a frame in $T_p^*\Sigma \otimes T_p^*\Sigma$ (tensor product of the dual frame). If h_{ab} are the components of h then we have the following representation of the generalized De Witt metric

$$G_{(\lambda)} = G_{(\lambda)}^{abcd}\,dh_{ab}\otimes dh_{cd}\,, \quad (157a)$$

where

$$G_{(\lambda)}^{abcd} = \frac{1}{2}\big(h^{ac}h^{bd} + h^{ad}h^{bc} - 2\lambda\,h^{ab}h^{cd}\big)\,. \quad (157b)$$

Here we introduced a factor λ in order to parametrize the impact of the negative trace term. We also consider Σ to be of general dimension n.

The inverse metric to (157) is given by

$$G_{(\lambda)}^{-1} = G_{(\lambda)\,abcd}^{-1}\frac{\partial}{\partial h_{ab}}\otimes\frac{\partial}{\partial h_{cd}}\,, \quad (158a)$$

where

$$G_{(\lambda)\,abcd}^{-1} = \frac{1}{2}\big(h_{ac}h_{bd} + h_{ad}h_{bc} - 2\mu\,h_{ab}h_{cd}\big)\,. \quad (158b)$$

The relation between λ and μ is

$$\lambda + \mu = n\lambda\mu\,, \quad (159)$$

so that

$$G_{(\lambda)}^{abnm}G_{(\lambda)\,nmcd}^{-1} = \tfrac{1}{2}\big(\delta^a_c\delta^b_d + \delta^a_d\delta^b_c\big)\,. \quad (160)$$

In ordinary GR $n = 3$, $\lambda = 1$, and $\mu = 1/2$. Note that there are good reasons the keep the superscript -1 even in component notation, that is, to write $G_{(\lambda)\,abcd}^{-1}$ rather than just $G_{(\lambda)\,abcd}$, since $G_{(\lambda)\,abcd}^{-1}$ does not equal $h_{ak}h_{bl}h_{cm}h_{dn}G_{(\lambda)}^{klmn}$ unless $\lambda = 2/n$, in which case $\lambda = \mu$.

If we change coordinates according to

$$\tau := \ln\Big(\big[\det\{h_{ab}\}\big]^{\frac{1}{n}}\Big)\,,$$
$$r_{ab} := h_{ab}/\big[\det\{h_{ab}\}\big]^{\frac{1}{n}}\,, \quad (161)$$

where τ parametrizes conformal changes and r_{ab} the conformally invariant ones, the metric (157) reads

$$G_{(\lambda)} = n(1-\lambda n)\,d\tau\otimes d\tau + r^{ac}r^{bd}\,dr_{ab}\otimes dr_{cd}\,, \quad (162)$$

where $r^{an}r_{nb} = \delta^a_b$. Since h is positive definite, so is r. Hence the second part is positive definite on the $\big(\frac{1}{2}n(n+1)-1\big)-$ dimensional vector space of trace-free symmetric tensors. Hence the De Witt metric is positive definite for $\lambda < 1/n$, Lorentzian for $\lambda > 1/n$, and simply degenerate (one-dimensional null space) for the critical value $\lambda = 1/n$. In the GR case we have $\lambda = 1$ and $n = 3$, so that the De Witt metric is Lorentzian of signature $(-,+,+,+,+,+)$. Note that this Lorentzian signature is independent of ε, i.e. it has nothing to do with the Lorentzian signature of the spacetime metric.

In the Hamiltonian formulation it is not G but rather a conformally related metric that is important, the conformal factor being $\sqrt{\det\{h_{ab}\}}$. If we set

$$\hat{G}_{(\lambda)} := \left[\det\{h_{ab}\}\right]^{1/2} G_{(\lambda)} \qquad (163)$$

and correspondingly

$$\hat{G}_{(\lambda)}^{-1} := \left[\det\{h_{ab}\}\right]^{-1/2} G_{(\lambda)}^{-1}, \qquad (164)$$

we can again write $\hat{G}_{(\lambda)}$ in terms of (τ, r_{ab}). In fact, the conformal rescaling clearly just corresponds to multiplying (162) with $\sqrt{\det\{h_{ab}\}} = e^{n\tau/2}$. Setting

$$T := 4\left[(1-n\lambda)/n\right]^{1/2} e^{n\tau/4} \qquad (165)$$

we get, excluding the degenerate case $\lambda = 1/n$,

$$\begin{aligned} \hat{G}_{(\lambda)} = \ &\mathrm{sign}(1 - n\lambda) \, dT \otimes dT \\ &+ T^2 \, C \, r^{ac} r^{bd} \, dr_{ab} \otimes dr_{cd}, \end{aligned} \qquad (166)$$

where $C = n/(16|1 - n\lambda|)$ ($= 3/32$ in GR). This is a simple warped product metric of $\mathbb{R}_+$ with the left-invariant metric on the homogeneous space $GL(3,\mathbb{R})/SO(3) \times \mathbb{R}_+$ of symmetric positive definite forms modulo overall scale, the warping function being just T^2 if T is the coordinate on $\mathbb{R}_+$. Now, generally, quadratic warped-product metrics of the form $\pm dT \otimes dT + T^2 g$, where g is independent of T, are non-singular for $T \searrow 0$ iff g is a metric of constant curvature ± 1 (like for a unit sphere in $\mathbb{R}^n$, with T being the radius coordinate, or the unit spacelike hyperboloid in n-dimensional Minkowski space, respectively). This is not the case for (166), which therefore has a curvature singularity for small T, i.e. small $\det\{h_{ab}\}$. Note that this is a singularity in the space of metrics (here at a fixed space point), which has nothing to do with spacetime singularities. In the early days of Canonical Quantum Gravity this has led to speculations concerning "natural" boundary conditions for the wave function, whose domain is the space of metrics [52]. The intention was to pose conditions such that the wave function should stay away from such singular regions in the space of metrics; see also [92] for a more recent discussion.

We stress once more that the signature of the De Witt metric is *not* related to the signature of spacetime, i.e. independent of ε. For example, for the GR values $\lambda = 1$ and $n = 3$, it is Lorentzian even if spacetime were given a Riemannian metric. Moreover, by integrating over Σ, the pointwise metric (166) defines a bilinear form on the infinite dimensional space of Riemannian structures on Σ, the geometry of which may be investigated to some limited extent [70][74].

7 Constrained Hamiltonian systems

In this section we wish to display some characteristic features of Hamiltonian dynamical systems with constraints. We restrict attention to finite-dimensional systems in order to not overload the discussion with analytical subtleties.

Let Q be the n-dimensional configuration manifold of a dynamical system that we locally coordinatize by $(q^1, \cdots, q^n)$. By TQ we denote its tangent bundle, which we coordinatize by $(q^1, \cdots, q^n, v^1, \cdots, v^n)$, so that a tangent vector $X \in TQ$ is given by $X = v^a \partial/\partial q^a$. The dynamics of the system is described by a *Lagrangian*

$$L : TQ \to \mathbb{R}, \qquad (167)$$

which selects the dynamically possible trajectories in TQ as follows: Let $\mathbb{R} \ni t \mapsto x(t) \in Q$ be a (at least twice continuously differentiable) curve, then it is dynamically possible iff the following *Euler Lagrange equations* hold (we set $dx/dt =: \dot{x}$):

$$\left.\frac{\partial L}{\partial q^a}\right|_{\substack{q=x(t) \\ v=\dot{x}(t)}} - \frac{d}{dt}\left[\left.\frac{\partial L}{\partial v^a}\right|_{\substack{q=x(t) \\ v=\dot{x}(t)}}\right] = 0. \qquad (168)$$

Performing the t-differentiation on the second term, this is equivalent to

$$H_{ab}\big(x(t), \dot{x}(t)\big)\ddot{x}^b = V_a\big(x(t), \dot{x}(t)\big), \qquad (169)$$

where

$$H_{ab}(q, v) := \frac{\partial^2 L(q, v)}{\partial v^a \partial v^b}, \qquad (170)$$

27

and

$$V_a(q,v) := \frac{\partial L(q,v)}{\partial q^a} - \frac{\partial^2 L(q,v)}{\partial v^a \partial q^b} \, v^b \,. \qquad (171)$$

Here we regard H and V as function on TQ with values in the symmetric $n \times n$ matrices and $\mathbb{R}^n$ respectively. In order to be able to solve (169) for the second derivative $\ddot{x}$ the matrix H has to be invertible, that is, it must have rank n. That is the case usually encountered in mechanics. On the other hand, *constrained systems* are those where the rank of H is not maximal. This is the case we are interested in.

We assume H to be of constant rank $r < n$. Then, for each point on TQ, there exist $s = (n - r)$ linearly independent kernel elements $K_{(\alpha)}(q,v)$, $\alpha = 1, \cdots, s$, such that $K^a_{(\alpha)}(q,v) H_{ab}(q,v) = 0$. Hence any solution $x(t)$ to (169) must be such that the curve $t \mapsto \big(x(t), \dot{x}(t)\big)$ in TQ stays on the subset

$$\mathcal{C} := \big\{ (q,v) \in TQ : \psi_\alpha(q,v) = 0 \,, \quad \alpha = 1, \cdots, s \big\}, \qquad (172a)$$

where

$$\psi_\alpha(q,v) = K^a_{(\alpha)}(q,v) V_a(q,v) \,. \qquad (172b)$$

We assume $\mathcal{C} \subset TQ$ to be a smooth closed submanifold of co-dimension s, i.e. of dimension $2n - s = n + r$.

Now we consider the cotangent bundle T^*Q over Q. On T^*Q we will use so-called *canonical coordinates*, denoted by $\{q^1, \cdots, q^n, p_1, \cdots, p_n\}$, the precise definition of which we will give below. The Lagrangian defines a map $\mathrm{FL} : TQ \to T^*Q$, which in these coordinates reads

$$\mathrm{FL}(q,v) = \left(q, p := \frac{\partial L(q,v)}{\partial v} \right) \,. \qquad (173)$$

From what has been said above it follows that the Jacobian of that map has constant rank $n + r$. Given sufficient regularity, we may further assume that

$$\mathcal{C}^* := \mathrm{FL}\big(\mathcal{C}\big) \subset T^*Q \qquad (174)$$

is a smoothly embedded closed submanifold in phase space T^*Q of co-dimension s. Hence there

are s functions ϕ_α, $\alpha = 1, \cdots, s$ such that

$$\mathcal{C}^* := \big\{ (q,p) \in T^*Q : \phi_\alpha(q,p) = 0 \,, \ \alpha = 1, \cdots, s \big\}. \qquad (175)$$

This is called the *constraint surface* in phase space. It is given as the intersection of the zero-level sets of s independent functions. Independence means that at each $p \in \mathcal{C}^*$ the s one-forms $d\phi_1|_p \cdots, d\phi_s|_p$ are linearly independent elements of $T^*_p T^*Q$.

The dynamical trajectories of our system will stay entirely on $\mathcal{C}^*$. The trajectories themselves are integral lines of a Hamiltonian flow. But what is the Hamiltonian function that generates this flow? To explain this we first recall the definition of the *energy function* for the Lagrangian L. It is a function $E : TQ \to \mathbb{R}$ defined through

$$E(q,v) := \frac{\partial L(q,v)}{\partial v^a} \, v^a - L(q,v) \,. \qquad (176)$$

At first sight this function cannot be defined on phase space, for we cannot invert FL to express v as function of q and p which we could insert into $E(q,v)$ in order to get $E(q, v(q,p))$. However, one may prove the following: There exists a function

$$H_{\mathcal{C}^*} : \mathcal{C}^* \to \mathbb{R} \,, \qquad (177a)$$

so that

$$E = H_{\mathcal{C}^*} \circ \mathrm{FL} \,. \qquad (177b)$$

A local version of this is seen directly from taking the differential of (176), which yields $dE = v^a d(\partial L / \partial v^a) - (\partial L / \partial q^a) dq^a$, expressing the fact that $dE_{(q,v)}(X) = 0$ if $\mathrm{FL}_{*(q,v)}(X) = 0$ for $X \in T_{(q,v)}TQ$, or in simple terms: E does not vary if q and p do not vary.

So far the function $H_{\mathcal{C}^*}$ is only defined on $\mathcal{C}^*$. By our regularity assumptions there exists a smooth extension of it to T^*Q, that is a function $H_0 : T^*Q \to \mathbb{R}$ such that $H_0|_{\mathcal{C}^*} = H_{\mathcal{C}^*}$. This is clearly not unique. But we can state the following: Let H_0 and H both be smooth (at least continuously differentiable) extensions of $H_{\mathcal{C}^*}$ to T^*Q, then there exist s smooth functions $\lambda^\alpha : T^*Q \to \mathbb{R}$ such that

$$H = H_0 + \lambda^\alpha \phi_\alpha \,. \qquad (178)$$

Locally a proof is simple: Let $f : T^*Q \to \mathbb{R}$ be continuously differentiable and such that $f|_{\mathcal{C}^*} \equiv 0$. Consider a point $p \in \mathcal{C}^*$ and coordinates $(x^1, \cdots, x^{2n-s}, y^1, \cdots y^s)$ in a neighborhood $U \subset T^*Q$ of p, where the x's are coordinates on the constraint surface and the y's are just the functions ϕ. In U the constraint surface is clearly just given by $y^1 = \cdots = y^s = 0$. Then

$$
\begin{aligned}
f|_U(x, y) &= \int_0^1 dt \frac{d}{dt} f(x, ty) \\
&= \int_0^1 dt \frac{\partial f}{\partial y^\alpha}(x, ty)\, y^\alpha = \lambda_\alpha(x, y)\, y^\alpha \,,
\end{aligned}
$$
$$(179a)$$

where

$$
\lambda_\alpha(x, y) := \int_0^1 dt \frac{\partial f}{\partial y^\alpha}(x, ty) \,. \tag{179b}
$$

For a global discussion see [80].

As *Hamiltonian* for our constraint system we address any smooth (at least continuously differentiable) extension H of $H_{\mathcal{C}^*}$. So if H_0 is a somehow given one, any other can be written as

$$
H = H_0 + \lambda^\alpha \phi_\alpha \tag{180}
$$

for some (at least continuously differentiable) real-valued functions λ^α on T^*Q.

Here we have been implicitly assuming that the Hamiltonian dynamics does not leave the constraint surface (174). If this were not the case we would have to restrict further to proper submanifolds of $\mathcal{C}^*$ such that the Hamiltonian vector fields evaluated on them lie tangentially. (If no such submanifold can be found the theory is simply empty). This is sometimes expressed by saying that the *primary constraints* (those encountered first in the Lagrangian/Hamiltonian analysis) are completed by *secondary, tertiary*, etc. constraints for consistency.

Here we assume that our system is already dynamically consistent. This entails that the Hamiltonian vector-fields X_{ϕ_α} for the ϕ_α are tangential to the constraint surface. This is equivalent to $X_{\phi_\alpha}(\phi_\beta)|_{\mathcal{C}^*} = 0$, or expressed in Poisson brackets:

$$
\{\phi_\alpha, \phi_\beta\}\big|_{\mathcal{C}^*} = 0 \,, \tag{181}
$$

for all $\alpha, \beta \in \{1, \cdots, s\}$. Following Dirac [55], constraints which satisfy this condition are said to be of *first class*. By the result shown (locally) above in (179) this is equivalent to the existence of $\frac{1}{2}s^2(s-1)$ (at least continuously differentiable) real-valued functions $C^\gamma_{\alpha\beta} = -C^\gamma_{\beta\alpha}$ on T^*Q, such that

$$
\{\phi_\alpha, \phi_\beta\} = C^\gamma_{\alpha\beta}\, \phi_\gamma \,. \tag{182}
$$

Note that as far as the intrinsic geometric properties of the constraint surface are concerned (181) and (182) are equivalent.

The indeterminacy of the Hamiltonian due to the freedom to choose any set of λ^α seems to imply an s-dimension worth of indeterminacy in the dynamically allowed motions. But the difference in these motions is that generated by the constraint functions on the constraint surface. In order to actually tell apart two such motions requires observables (phase-space functions) whose Poisson brackets with the constraints do *not* vanish on the constraint surface. The general attitude is to assume that this is not possible, i.e. to assume that *physical observables* correspond exclusively to phase-space functions whose Poisson bracket with all constraints vanish on the constraint surface. This is expressed by saying that all motions generated by the constraints are *gauge transformations*. This entails that they are undetectable in principle and merely correspond to a mathematical redundancy in the description rather than to any physical degrees of freedom. It is therefore more correct to speak of gauge *redundancies* rather than of gauge *symmetries*, as it is sometimes done, for the word "symmetry" is usually used for a physically meaningful operation that does change the object to which it is applied in at least some aspects (otherwise the operation is the identity). Only some "relevant" aspects, in the context of which one speaks of symmetry, are not changed.

7.1 Geometric theory

Being first class has an interpretation in terms of symplectic geometry. To see this, we first recall a few facts and notation from elementary symplec-

tic geometry of cotangent bundles. Here some sign conventions enter and the reader is advised to compare carefully with other texts.

A *symplectic structure* on a manifold is a non-degenerate closed two-form. Such structures always exist in a natural way on cotangent bundles, where they even derive from a *symplectic potential*. The latter is a one-form field θ on T^*Q whose general geometric definition is as follows: Let $\pi : T^*Q \to Q$ be the natural projection from the co-tangent bundle of Q (phase space) to Q itself. Then, for each $p \in T^*Q$, we define

$$\theta_p := p \circ \pi_{*p} \,. \tag{183}$$

So in order to apply θ_p to a vector $X \in T_pT^*Q$, we do the following: Take the differential π_* of the projection map π, evaluate it at point p and apply it to $X \in T_pT^*Q$ in order to push it forward to the tangent space $T_{\pi(p)}Q$ at point $\pi(p) \in Q$. Then apply p to it, which makes sense since p is, by definition, an element of the co-tangent space at $\pi(p) \in Q$.

The symplectic structure, ω, is now given by

$$\omega = -\,d\theta \,. \tag{184}$$

The minus sign on the right-hand side has no significance other than to comply with standard conventions. Let us stress that θ, and hence ω, is globally defined. This is obvious from the global definition (183). Therefore ω is not only closed, $d\omega = 0$, but even globally exact for any Q. Non-degeneracy of ω will be immediate from the expression in canonical coordinates to be discussed below (cf. (196b)).

A diffeomorphism $F : T^*Q \to T^*Q$ is called a *canonical transformation* or *symplectic morphism* if it preserves ω, that is, if $F^*\omega = \omega$. We explicitly mention two kinds of canonical transformations, which in some sense are complementary to each other.

The first set of canonical transformations are fibre-preserving ones. This means that, for each $q \in Q$, points in the fibre $\pi^{-1}(q)$ are moved to points in the same fibre $\pi^{-1}(q)$. This is equivalent to the simple equation

$$\pi \circ F = \pi \,. \tag{185}$$

The special fibre-preserving diffeomorphisms we wish to mention are given by adding to each momentum $p \in T^*Q$ the value $\sigma\big(\pi(p)\big)$ of a section $\sigma : Q \to T^*Q$:

$$F(p) = p + \sigma_{\pi(p)} \,. \tag{186}$$

This transforms the symplectic potential at $p \in T^*Q$ into

$$
\begin{aligned}
(F^*\theta)_p &= \theta_{F(p)} \circ F_{*p} \\
&\overset{(183)}{=} F(p) \circ \big(\pi \circ F\big)_{*p} \\
&\overset{(185)}{=} F(p) \circ \pi_{*p} \\
&\overset{(186)}{=} \theta_p + \sigma_{\pi(p)} \circ \pi_{*p} \\
&= \theta_p + \big(\pi^*\sigma\big)_p \,.
\end{aligned}
\tag{187}
$$

Hence

$$
\begin{aligned}
F^*\theta &= \theta + \pi^*\sigma \,, &\tag{188a} \\
F^*\omega &= \omega - \pi^*d\sigma \,. &\tag{188b}
\end{aligned}
$$

This is a canonical transformation if σ is a closed covector field on Q. By Poincaré's Lemma such a σ is locally exact, but this need not be the case globally. Obstructions to global exactness are the first De Rahm cohomology class $H^1_{\mathrm{DR}}(Q)$, which is just defined to be the vector space of closed modulo exact covector fields on Q. The dimension of this vector space equals the rank of the free part of the ordinary first homology group $H_1(Q, \mathbb{Z})$ on Q with integer coefficients. This latter group is always abelian and isomorphic to the abelianization of the (generally non-abelian) first homotopy group $\pi_1(Q)$. Hence for non-simply connected Q the possibility of canonical transformations exist which change the symplectic potential by a closed yet non-exact covector field.

The second set of canonical transformations that we wish to mention are natural extensions to T^*Q of diffeomorphisms of Q. These extensions not only leave invariant the symplectic structure ω but also the symplectic potential θ. To see this we note that any diffeomorphism $f : Q \to Q$ has a natural *lift* to T^*Q. We recall that a *lift* of a diffeomorphism

f of the base manifold Q is a diffeomorphism $F :$ $T^*Q \to T^*Q$ such that

$$\pi \circ F = f \circ \pi \, . \tag{189}$$

This is equivalent to saying that the following diagram of maps commutes (a tailed arrow indicates injectivity and a double-headed arrow surjectivity)

$$\tag{190}$$

Here the map F is just the pull-back of the inverse f^{-1}. Hence the image of $p \in T^*Q$ is given by

$$F(p) = p \circ f^{-1}_{*f(\pi(p))} \, . \tag{191}$$

From that it follows that the symplectic potential is invariant under all lifts of diffeomorphisms on Q:

$$
\begin{aligned}
(F^*\theta)_p &= \theta_{F(p)} \circ F_{*p} \\
&\overset{(183)}{=} F(p) \circ \big(\pi \circ F\big)_{*p} \\
&\overset{(189)}{=} F(p) \circ f_{*\pi(p)} \circ \pi_{*p} \\
&\overset{(191)}{=} p \circ f^{-1}_{*f(\pi(p))} \circ f_{*\pi(p)} \circ \pi_{*p} \\
&= p \circ \big(f^{-1} \circ f\big)_{*\pi(p)} \circ \pi_{*p} \\
&= \theta_p \, .
\end{aligned}
\tag{192}
$$

So far we deliberately avoided intoducing local coordinates in order to stress global existence of the quantities in question. We now introduce convenient coordinates in which the symplectic potential and structure take on the familiar form. These are called *canonical coordinates*, which we already mentioned above and the definition of which we now give. Let (x, U) be a local chart on Q such that $x : Q \supset U \to \mathbb{R}^n$ is the chart map with component functions x^a. This chart induces a chart (z, V) on T^*Q, where $V = \pi^{-1}(U) \subset T^*Q$ and $z : V \to \mathbb{R}^{2n}$. We follow general tradition and label

the first set of n component functions by $z^a = q^a$ (for $a = 1, \cdots, n$) and the second set by $z^{n+a} = p_a$ (for $a = 1, \cdots, n$). For the first set we define

$$q^a(\lambda) := x^a\big(\pi(\lambda)\big) \, , \tag{193a}$$

and for the second

$$p_a(\lambda) := \lambda\left(\left.\frac{\partial}{\partial x^a}\right|_{\pi(\lambda)}\right) \, , \tag{193b}$$

for any $\lambda \in V$. Note that (193b) just says that $\lambda = p_a(\lambda)\, dx^a|_{\pi(\lambda)}$. In this way we get a "canonical" extension of any chart on Q with domain U to a chart on T^*Q with domain $V = \pi^{-1}(U)$. From the definition it is clear that

$$\pi_{*\lambda}\left(\left.\frac{\partial}{\partial q^a}\right|_\lambda\right) = \left.\frac{\partial}{\partial x^a}\right|_{\pi(\lambda)} \tag{194a}$$

and

$$\pi_{*\lambda}\left(\left.\frac{\partial}{\partial p_a}\right|_\lambda\right) = 0 \, . \tag{194b}$$

It immediately follows from the definition (183) that

$$\theta_\lambda\left(\left.\frac{\partial}{\partial q^a}\right|_\lambda\right) = p_a(\lambda) \tag{195a}$$

and

$$\theta_\lambda\left(\left.\frac{\partial}{\partial p_a}\right|_\lambda\right) = 0 \, . \tag{195b}$$

Hence, in canonical coordinates, the symplectic potential and structure take on the form

$$\theta|_V = p_a \, dq^a \, , \tag{196a}$$
$$\omega|_V = dq^a \wedge dp_a \, . \tag{196b}$$

Note again that (196) is valid in any canonical completion of a chart on Q. As advertised above, it is immediate from (196b) that $\omega|_V$ is non-degenerate at any point $p \in V$. Since non-degeneracy is a pointwise property and valid in any canonical chart, it follows that ω is non-degenerate everywhere. In the sequel we shall drop the explicit mention of the chart domain V.

The non-degeneracy of ω allows to uniquely associate a vector field X_f to any real-valued function f on T^*Q through

$$i_{X_f}\omega = df \, . \tag{197}$$

It is called the *Hamiltonian vector field* of f. An immediate consequence of (197) and $d\omega = 0$ is that ω has vanishing Lie derivative with respect to any Hamiltonian vector field:

$$L_{X_f}\omega = (i_{X_f} \circ d + d \circ i_{X_f})\omega = 0 \, . \qquad (198)$$

In coordinates X_f looks like this:

$$X_f = \frac{\partial f}{\partial p_a}\frac{\partial}{\partial q^a} - \frac{\partial f}{\partial q^a}\frac{\partial}{\partial p_a} \, . \qquad (199)$$

The *Poisson bracket* between two functions f and g is defined as

$$\{f,g\} := \omega(X_f, X_g) = X_g(f) = -X_f(g) \quad (200a)$$
$$= \frac{\partial f}{\partial q^a}\frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q^a} \, . \qquad (200b)$$

It provides $C^\infty(T^*Q)$ with a structure of a Lie algebra, which means that for all $f, g, h \in C^\infty(T^*Q)$ and all $a \in \mathbb{R}$ we have,

$$\{f,g\} = -\{g,f\} \, , \qquad (201a)$$
$$\{af + g, h\} = a\{f,h\} + \{g,h\} \, , \qquad (201b)$$
$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0 \, . \qquad (201c)$$

Antisymmetry and bi-linearity are obvious from (200). The third property (201c), called the *Jacobi identity*, can of course be directly checked using the coordinate expression (200b), but the geometric proof is more instructive, which we therefore wish to present here.

The first thing we note is that the map $f \mapsto X_f$ obeys

$$X_{\{f,g\}} = -[X_f, X_g] \, . \qquad (202)$$

This follows from

$$\begin{aligned}
d\{f,g\} &= d\big(\omega(X_f, X_g)\big) \\
&= d i_{X_g} i_{X_f}\omega \\
&= L_{X_g} i_{X_f}\omega - i_{X_g} d i_{X_f}\omega \\
&= i_{[X_g, X_f]}\omega = -i_{[X_f, X_g]}\omega \, ,
\end{aligned} \qquad (203)$$

where in the last step we used $d i_{X_f}\omega = 0$ due to $d\omega = 0$ and $L_{X_f}\omega = 0$ and once more $L_{X_g}\omega = 0$ in

the first term. As ω is non-degenerate comparison with (197) leads to (202). Next we recall that the exterior differential of a general k-form field α, applied to the $k+1$ vectors $X_0, X_1, \cdots, X_k$, can be written as

$$\begin{aligned}
&d\alpha(X_0, \cdots, X_k) \\
&= \sum_{0 \le i \le k} (-1)^i \, X_i\big(\alpha(X_0, \cdots, \hat{X}_i, \cdots, X_k)\big) \\
&+ \sum_{0 \le i < j \le k} (-1)^{i+j} \, \alpha\big([X_i, X_j], X_0, \cdots, \\
&\qquad\qquad\qquad \hat{X}_i, \cdots, \hat{X}_j, \cdots, X_k\big) \, .
\end{aligned} \qquad (204)$$

Here the hatted entries are omitted. This we apply to $d\omega(X_f, X_g, X_h)$ and make use of, e.g., $X_f\big(\omega(X_g, X_h)\big) = (L_{X_f}\omega)(X_g, X_h) + \omega([X_f, X_g], X_h) + \omega(X_g, [X_f, X_h])$, as well as (198). Then we get, writing $\sum_{(fgh)}$ for the cyclig sum over f, g, h:

$$\begin{aligned}
d\omega(X_f, X_g, X_h) &= -\sum_{(fgh)} \omega\big(X_f, [X_g, X_h]\big) \\
&\overset{(202)}{=} \sum_{(fgh)} \omega\big(X_f, X_{\{g,h\}}\big) \\
&\overset{(200a)}{=} \sum_{(fgh)} \{f, \{g,h\}\} \, .
\end{aligned} \qquad (205)$$

Hence we see that the Jacobi identity (201c) follows from $d\omega = 0$.

Equations (201) state that $C^\infty(T^*Q)$ is a Lie algebra with $\{\cdot, \cdot\}$ as Lie product. With respect to the ordinary pointwise product of functions $C^\infty(T^*Q)$ is already a commutative and associative algebra. Both structures are linked via

$$\{f \cdot g, h\} = f \cdot \{g,h\} + \{f,h\} \cdot g \, , \qquad (206)$$

which immediately follows from the Leibniz rule and (200) (which implies that $\{f \cdot g, h\}$ is the derivative of $f \cdot g$ along X_h). All this is expressed by saying that $C^\infty(T^*Q)$ is a *Poisson algebra*, meaning the coexistence of two algebraic structures, one of a commutative and associative algebra, one of a Lie algebra, and their compatibility via (206). Note that the liner space $\Gamma T(T^*Q)$ of vector fields

is also a Lie algebra, whose Lie product is the commutator of vector fields. Equation (202) then shows that the map from $C^\infty(T^*Q)$ to $\Gamma T(T^*Q)$ that sends $f \mapsto X_f$ is an anti homomorphism of Lie algebras.

After this brief digression we now return to the geometric interpretation of first-class constraints. For any $p \in \mathcal{C}^*$ we define

$$T_p^\perp(T^*Q) := \\ \left\{ X \in T_p(T^*Q) : \omega(X,Y) = 0 \,, \forall Y \in T_p\mathcal{C}^* \right\}. \tag{207}$$

The non-degeneracy of ω implies that the dimension of $T_p^\perp(T^*Q)$ equals s, the co-dimension of $\mathcal{C}^*$ in T^*Q. But note that as ω is skew, $T_p^\perp(T^*Q)$ might well have a non-trivial intersection with $T_p\mathcal{C}^*$. This gives rise to the following characterizations for the submanifold $\mathcal{C}^* \subset T^*Q$ (understood to hold at each point $p \in \mathcal{C}^*$): $\mathcal{C}^*$ is called

- *isotropic* iff $T_p\mathcal{C}^* \subset T_p^\perp(T^*Q)$;

- *co-isotropic* iff $T_p\mathcal{C}^* \supset T_p^\perp(T^*Q)$;

- *Lagrangian* iff $T_p\mathcal{C}^* = T_p^\perp(T^*Q)$.

Since $\{\phi_\alpha, \phi_\beta\} = d\phi_\alpha(X_{\phi_\beta})$ we see that condition (181) is equivalent to the statement that the Hamiltonian vector-fields for the constraint functions ϕ_α are tangent to the constraint hypersurface:

$$X_{\phi_\alpha}|_{\mathcal{C}^*} \in \Gamma T\mathcal{C}^*. \tag{208}$$

Our assumption that the s differentials $d\phi_\alpha$ be linearly independent at each $p \in \mathcal{C}^*$ now implies that the s vectors $X_{\phi_\alpha}(p)$ span an s-dimensional subspace of $T_p\mathcal{C}^*$. But they are also elements of $T_p^\perp(T^*Q)$ since $\omega(X_{\phi_\alpha}, Y) = d\phi_\alpha(Y) = 0$ for all Y tangent to $\mathcal{C}^*$. As the dimension of $T_p^\perp(T^*Q)$ is s, this shows

$$T_p^\perp(T^*Q) = \text{span}\{X_{\phi_1} \cdots, X_{\phi_s}\} \subset T_p\mathcal{C}^*, \tag{209}$$

that is, co-isotropy of $\mathcal{C}^*$. First-class constraints are precisely those which give rise to co-isotropic constraint surfaces.

The significance of this lies in the following result, which we state in an entirely intrinsic geometric fashion. Let $\mathcal{C}^* \subset T^*Q$ be co-isotropic of co-dimension s and let $e : \mathcal{C}^* \to T^*Q$ be its embedding. We write

$$\hat{\omega} := e^*\omega \tag{210}$$

for the pull back of ω to the constraint surface (i.e. essentially the restriction of ω to the tangent bundle of the constraint surface). $\hat{\omega}$ is now s-fold degenerate, its kernel at $p \in \mathcal{C}^*$ being just $T_p^\perp(T^*Q) \subset T_p\mathcal{C}^*$. We have the smooth assignment of subspaces

$$\mathcal{C}^* \ni p \mapsto \text{kernel}_p(\hat{\omega}) = T_p^\perp(T^*Q), \tag{211}$$

which forms a sub-bundle of $T\mathcal{C}^*$ called the *kernel distribution* of $\hat{\omega}$. Now, the crucial result is that this sub-bundle is *integrable*, i.e tangent to locally embedded submanifolds $\gamma^* \subset \mathcal{C}^*$ of co-dimension s in $\mathcal{C}^*$, or co-dimension $2s$ in T^*Q. Indeed, in order to show this we only need to show that whenever two vector fields X and Y on $\mathcal{C}^*$ take values in the kernel distribution their commutator $[X, Y]$ also takes values in the kernel distribution. That this suffices for local integrability is known as *Frobenius' theorem* in differential geometry. Writing

$$i_{[X,Y]}\hat{\omega} = L_X\left(i_Y\hat{\omega}\right) - i_Y\left(L_X\hat{\omega}\right) \tag{212}$$

we infer that the first term on the right-hand side vanishes because Y is in $\hat{\omega}$'s kernel and $L_X\hat{\omega}$ vanishes because $L_X = d \circ i_X + i_X \circ d$ on forms, where $i_X\hat{\omega} = 0$ again due to X being in the kernel and $d\hat{\omega} = de^*\omega = e^*d\omega = 0$ due to ω being closed.

The program of *symplectic reduction* is now to form the $(2n - 2s)$-dimensional quotient space $\mathcal{C}^*/\sim$, where $\sim$ is the equivalence relation whose equivalence classes are the maximal integral submanifolds of the kernel distribution of $\hat{\omega}$. $\mathcal{C}^*/\sim$ is called the *physical phase-space* or *reduced space of states*.

We stress that this geometric formulation of the reduction program, and the characterization of $\mathcal{C}^*/\sim$ in particular, does not refer to any set of functions ϕ_α that one might use in order to characterize $\mathcal{C}^*$. If one uses such functions, it is understood

that they obey the above-mentioned regularity conditions of being at least continuously differentiable in a neighborhood of $\mathcal{C}^*$ and giving rise to a set of s linearly independent differentials $d\phi_\alpha$ at any point of $\mathcal{C}^*$. Hence redefinitions of constraint functions like $\phi \mapsto \sqrt{|\phi|}$ or $\phi \mapsto \phi^2$, albeit leading to the same surface $\mathcal{C}^*$, are a priori not allowed.

The reduced phase space can be identified with the set of physical states. Smooth functions on this space then correspond to physical observables. But how can we characterize the latter without explicitly constructing $\mathcal{C}^*/\sim$? This we shall explain in the remaining part of this section.

We define the *gauge algebra* as the set of smooth functions on unreduced phase space that vanish on the constraint surface:

$$\mathrm{Gau} := \left\{ f \in C^\infty(T^*Q) : f|_{\mathcal{C}^*} = 0 \right\}. \qquad (213)$$

This set is clearly an associative ideal with respect to the pointwise product. But it is also a Lie subalgebra with respect to the Poisson bracket. To see this, we first remark that $X_f|_{\mathcal{C}^*}$ is $T^\perp(T^*Q)$-valued if $f \in \mathrm{Gau}$. Indeed, $f \in \mathrm{Gau}$ implies that $\mathrm{kernel}(df_p)$ includes $T_p\mathcal{C}^*$ for $p \in \mathcal{C}^*$. But then (197) shows $X_f(p) \in T_p^\perp(T^*Q)$. Now, this immediately implies that Gau is a Lie algebra for co-isotropic $\mathcal{C}^*$, for then (200) implies $\{f,g\}|_{\mathcal{C}^*} = \omega(X_f, X_g)|_{\mathcal{C}^*} = 0$. Hence Gau is a Poisson subalgebra of $C^\infty(T^*Q)$ and also an associative ideal with respect to pointwise multiplication. However, it is *not* a Lie-ideal with respect to $\{\cdot,\cdot\}$. Indeed, if $f \in \mathrm{Gau}$ we just need to take a $g \in C^\infty(T^*Q)$ which is not constant along the flow of $X_f|_{\mathcal{C}^*}$; then $\{f,g\}|_{\mathcal{C}^*} \neq 0$. This means that we cannot define physical observables by the quotient $C^\infty(T^*Q)/\mathrm{Gau}$, since this will not result in a Poisson algebra. As we insist that all elements in Gau generate gauge transformations, we have no choice but to reduce the size of $C^\infty(T^*Q)$ in order to render the quotient a Poisson algebra. Economically the most effective possibility is to take the Lie-idealizer of Gau in $C^\infty(T^*Q)$, which is defined as follows

$$I_{\mathrm{Gau}} := \left\{ f \in C^\infty(T^*Q) : \{f,g\}\big|_{\mathcal{C}^*} = 0 \ \forall g \in \mathrm{Gau} \right\}. \qquad (214)$$

Note that I_{Gau} is the set of smooth functions that, to use a terminology introduced by Dirac[55], *weakly* (Poisson) commute with the constraints (i.e. with Gau). Here *weak (Poisson) commutativity* means that the Poisson brackets of observables and constraints need not vanish globally, i.e. on T^*Q, but only after restriction to $\mathcal{C}^*$. We will briefly come back to the case of *strong (Poisson) commutativity* below.

Now, if $I_{\mathrm{Gau}}/\mathrm{Gau}$ is to make sense as Poisson algebra of physical observables I_{Gau} must be a Poisson algebra containing Gau as an Poisson ideal.

That I_{Gau} is an associative algebra under pointwise multiplication immediately follows from (206). That it is also a Lie subalgebra follows from the Jacobi identity (201c). Indeed, let $f, g \in I_{\mathrm{Gau}}$ and $h \in \mathrm{Gau}$; then (201c) immediately gives

$$\{\{f,g\},h\}\big|_{\mathcal{C}^*}$$
$$= -\underbrace{\{\{g,h\},f\}}_{\in\mathrm{Gau}}\big|_{\mathcal{C}^*} - \underbrace{\{\{h,f\},g\}}_{\in\mathrm{Gau}}\big|_{\mathcal{C}^*} = 0. \qquad (215)$$

Hence we have shown that I_{Gau} is a Poisson subalgebra of $C^\infty(T^*M)$ which contains Gau as a Poisson ideal. By construction I_{Gau} is the largest Poisson subalgebra of $C^\infty(T^*M)$ with that property. Hence we may identify the Poisson algebra of physical observables, or *reduced space of observables* with the quotient

$$\mathcal{O}_{\mathrm{phys}} := I_{\mathrm{Gau}}/\mathrm{Gau}. \qquad (216)$$

This complements the definition of the reduced space of physical states. Again we stress that the definition given here does not refer to any set of functions ϕ_α that one might use in order to characterize $\mathcal{C}^*$.

We also stress that instead of the Lie idealizer (214) we could *not* have taken the Lie centralizer

$$C_{\mathrm{Gau}} := \{ f \in C^\infty(T^*Q) : \{f,g\} = 0 \ \forall g \in \mathrm{Gau} \}. \qquad (217)$$

Note that here the only difference to (214) is that $\{f,g\}$ is required to vanish strongly (i.e. on all of T^*Q) and not only weakly (i.e. merely on $\mathcal{C}^*$). This makes a big difference and the quotient $I_{\mathrm{Gau}}/\mathrm{Gau}$ will now generally be far too small. In fact, it is

intuitively clear and also easy to prove (see, e.g., Lemma 5 of [72]) that the Hamiltonian vector fields X_f corresponding to functions $f \in$ Gau, which span $T_p^\perp(T^*\mathcal{C}) \subset T_p\mathcal{C}^*$ at each $p \in \mathcal{C}$, span all of $T_p(T^*Q)$ for each point p off $\mathcal{C}^*$. This implies that smooth functions in $C^\infty(T^*Q)$ which strongly (Poisson) commute with Gau are locally constant outside $\mathcal{C}^*$. Sometimes strong (Poisson) commutativity is required not with respect to Gau but with respect to a complete set $\{\phi_1, \cdots, \phi_s\}$ of functions in Gau defining $\mathcal{C}^*$; for example the component functions of the momentum map (see next subsection). But even then strong commutativity is too strong, as a smooth function commuting with all ϕ_α on $\mathcal{C}^*$ need generally not extend to a smooth function defined in a neighborhood of $\mathcal{C}^*$ in T^*Q which still commutes with all ϕ_α. The reason is that the leaves of the foliation defined by the ϕ_α may become 'wild' off $\mathcal{C}^*$. Compare, e.g., the discussion in [35] (including the example on p. 116).

7.2 First-class constraints from zero momentum-maps

First class constraints often arise from group actions (see Appendix for group actions). This is also true in GR, at least partially. So let us explain this in more detail. Let a Lie group G act on the left on T^*Q. This means that there is a map $G \times T^*Q \to T^*Q$, here denoted simply by $(g,p) \mapsto g \cdot p$, so that $g_1 \cdot (g_2 \cdot p) = (g_1 g_2) \cdot p$ and $e \cdot p = p$ if $e \in G$ is the neutral element. As already seen earlier in (24) and explained in detail in the Appendix, there is then an anti-homomorphism from $\mathrm{Lie}(G)$, the Lie algebra of G, to the Lie algebra of vector fields on T^*Q. Recall that the vector field V^X corresponding to $X \in \mathrm{Lie}(G)$, evaluated at point $p \in T^*Q$, is given by

$$V^X(p) := \frac{d}{dt}\Big|_{t=0} \exp(tX) \cdot p. \qquad (218)$$

Then

$$\left[V^X, V^Y\right] = -V^{[X,Y]}. \qquad (219)$$

Let us further suppose that the group action on T^*Q is of a special type, namely it arises from a group action on Q by a canonical lift. (Every diffeomorphism f of Q can be lifted to a diffeomorphism F of T^*Q given by the pull back of the inverse f^{-1}.) Then it is easy to see from the geometric definition (183) that the symplectic potential θ is invariant under this group action and consequently the group acts by symplectomorphisms (ω preserving diffeomorphisms). The infinitesimal version of this statement is that, for all $X \in \mathrm{Lie}(G)$,

$$L_{V^X}\theta = 0. \qquad (220)$$

Since $L_{V^X} = i_{V^X} \circ d + d \circ i_{V^X}$ this is equivalent to

$$i_{V^X}\omega = d\big(\theta(V^X)\big) \qquad (221)$$

which says that V^X is the Hamiltonian vector field of the function $\theta(V^X)$. We call the map

$$\mathrm{Lie}(G) \ni X \mapsto P(X) := \theta(V^X) \in C^\infty(T^*Q) \qquad (222)$$

the *momentum map* for the action of G. It is a linear map from $\mathrm{Lie}(G)$ to $C^\infty(T^*Q)$ and satisfies

$$\begin{aligned}
\big\{P(X), P(Y)\big\} &= V^Y\big(\theta(V^X)\big) \\
&= \big(L_{V^Y}\theta\big)(V^X) + \theta\big(L_{V^Y}V^X\big) \\
&= \theta\big(V^{[X,Y]}\big) \\
&= P\big([X,Y]\big),
\end{aligned} \qquad (223)$$

where we used (220) and (219) for the third equality. Hence we see that the map (222) is a Lie homomorphism from $\mathrm{Lie}(G)$ into the Lie algebra of smooth, real-valued functions on T^*Q (whose Lie product is the Poisson bracket).

Now, first class constraints are often given by the condition of *zero momentum mappings*, i.e., by $P(X) = 0$ for all $X \in \mathrm{Lie}(G)$. By linearity in X, this is equivalent to the set of $s := \dim(G)$ conditions

$$\phi_\alpha := P(e_\alpha) = 0, \qquad (224)$$

where $e_\alpha = \{e_1, \cdots, e_s\}$ is a basis of $\mathrm{Lie}(G)$. Let the structure constants for this basis be $C^\gamma_{\alpha\beta}$, i.e. $[e_\alpha, e_\beta] = C^\gamma_{\alpha\beta}e_\gamma$, then (223) becomes

$$\{\phi_\alpha, \phi_\beta\} = C^\gamma_{\alpha\beta}\phi_\gamma. \qquad (225)$$

Constraints in gauge theories will typically arise as zero momentum maps in the fashion described here, the only necessary generalization being the extension to infinite-dimensional groups and Lie algebras. In fact, for gauge theories our G will correspond to the infinite-dimensional *group of gauge transformations*, which is not to be confused with the finite-dimensional gauge group. The former consists of functions, or sections in bundles, with values in the latter. On the other hand, the constraints in GR will only partially be of this type. More precisely, those constraints arising from 3-dimensional diffeomorphisms (called the vector or diffeomorphism constraints) will be of this type, those from non-tangential hypersurface deformations (scalar or Hamiltonian constraint) will not fit into this picture. For the former G will correspond to $\text{Diff}(\Sigma)$, or some appropriate subgroup thereof, and $\text{Lie}\big(\text{Diff}(\Sigma)\big)$ to the infinite-dimensional Lie algebra of vector fields on Σ (possibly with special support and/or fall-off conditions). The different nature of the latter constraint will be signaled by structure *functions* $C_{\alpha\beta}^{\gamma}(q,p)$ appearing on the right-hand side rather than constants. This has recently given rise to attempts to generalize the group-theoretic setting described above to that of *groupoids* and *Lie algebroids*, in which the more general structure of GR can be accommodated [33].

8 Hamiltonian GR

The Hamiltonian formulation of GR proceed along the lines outlined in the previous section. For this we write down the action in a (3+1)-split form, read off the Lagrangian density, define the conjugate momenta as derivatives of the latter with respect to the velocities, and finally express the energy function (176) in terms of momenta. The constraint functions will not be determined on the Lagrangian level, but rather directly on the Hamiltonian level as primary and secondary constraints (there will be no tertiary ones), the primary ones being just the vanishing of the momenta for lapse and shift.

The Lagrangian density for GR is essentially just the scalar curvature of spacetime. However, upon variation of this quantity, which contains second derivatives in the metric, we will pick up boundary terms from partial integrations which need not vanish by just keeping the metric on the boundary fixed. Hence we will need to subtract these boundary terms which will otherwise obstruct functional differentiability. Note that this is not just a matter of aesthetics: Solutions to differential equations (like Einstein's equation) will not be stationary points of the action if the latter is not differentiable at these points. Typically, Euler Lagrange equations will allow for solutions outside the domain of differentiability of the action they are derived from. Including some such solutions will generally need to adapt the action by boundary terms. This clearly matters if one is interested in the values of the action, energies, etc. for these solutions and also, of course, in the path-integral formulations of the corresponding quantum theories.

The Einstein-Hilbert action of GR is

$$S_{\text{GR}}[\Omega,g] = -\frac{\varepsilon}{2\kappa}\int_{\Omega}\mathbf{Scal}\,d\mu_g + \text{boundary terms} \tag{226}$$

where in local coordinates $x^{\mu} = (x^0 = ct, x^1, x^2, x^3)$,

$$d\mu_g = \sqrt{\varepsilon\det\{g_{\mu\nu}\}}\,cdt\wedge dx^1\wedge dx^2\wedge dx^3\,. \tag{227}$$

The sign convention behind the prefactor $-\varepsilon$ in (226) is such that in the Lorentzian as well as the Riemannian case the Lagrangian density contains the bilinear De Witt inner product of the extrinsic curvatures (compare (141)) with a positive sign, i.e. transverse traceless modes have positive kinetic energy.

The boundary term can be read off (141) and (139). If the integration domain $\Omega \subset M$ is such that the spacelike boundaries are contained in two hypersurfaces Σ_s, i.e. two $t = \text{const.}$ surfaces, say $\partial\Omega_i := \partial\Omega\cap\Sigma_{\text{initial}}$ and $\partial\Omega_f := \partial\Omega\cap\Sigma_{\text{final}}$, we would have to add the two boundary terms (dependence on ε drops out)

$$\kappa^{-1}\int_{\partial\Omega_f}\text{Tr}_h(K)\,d\mu_h - \kappa^{-1}\int_{\partial\Omega_i}\text{Tr}_h(K)\,d\mu_h\,. \tag{228}$$

Here we used that the second term in (139) does not contribute due to a being orthogonal to n. $d\mu_h$ is the standard measure from the induced metric, h, on the hypersurfaces. If the cylindrical timelike boundary $\partial\Omega_{\text{cyl}}$ is chosen such that its spacelike normal m is orthogonal to n, only the second term in (139) contributes and we get one more boundary term (again ε drops out)

$$\kappa^{-1} \int_{\partial\Omega_{\text{cyl}}} \hat{K}(n,n)\, d\mu_h \,. \qquad (229)$$

Here $\hat{K}$ is the extrinsic curvature of $\partial\Omega_{\text{cyl}}$ in M, which we picked up because $g(a,m) = g(\nabla_n n, m) = -g(n, \nabla_n m) = \hat{K}(n,n)$.

Once the boundary terms are taken care of we can just read off the Lagrangian density from (141) also using (66),

$$\mathcal{L}_{\text{GR}} = (2\kappa)^{-1} \big[G(K,K) - \varepsilon R \big] \alpha \sqrt{h} \,, \qquad (230)$$

where we now used the standard abbreviations

$$\begin{aligned}
G(K,K) & := G^{abcd} K_{ab} K_{cd} \,, \\
R & := \mathbf{Scal}^D \,, \qquad (231) \\
\sqrt{h} & := \sqrt{\det\{h_{ab}\}} \,.
\end{aligned}$$

Moreover, K_{ab} has here to be understood as expressed in terms of the time and Lie derivatives of h_{ab}:

$$K = -\tfrac{\varepsilon}{2}\alpha^{-1}\big(\dot{h} - L_\beta h\big)\,. \qquad (232)$$

We keep in mind that an overdot denotes differentiation with respect to ct (not t). In passing we also note that $\mathcal{L}_{\text{GR}}$ has the right physical dimension of an energy-density (α is dimensionless).

The Hamiltonian density is now obtained by the usual Legendre transform with respect to all configuration variables that are varied in the action. These comprise all components $g_{\mu\nu}$ and hence in the (3+1)-split parametrization all h_{ab} as well as the lapse α and the three shift components β^a. However, it is immediate that (230) does not contain any time derivatives of the latter; hence their

conjugate momenta vanish:

$$\pi_\alpha := \frac{1}{c}\frac{\partial\mathcal{L}_{\text{GR}}}{\partial\dot{\alpha}} = 0\,, \qquad (233\text{a})$$

$$\pi_{\beta^a} := \frac{1}{c}\frac{\partial\mathcal{L}_{\text{GR}}}{\partial\dot{\beta}^a} = 0\,. \qquad (233\text{b})$$

This leaves us with the momenta for the metric components h_{ab}

$$\begin{aligned}
\pi^{ab} & := \frac{1}{c}\frac{\partial\mathcal{L}_{\text{GR}}}{\partial\dot{h}_{ab}} \\
& = \frac{(-\varepsilon)\sqrt{h}}{2\kappa c} G^{abcd} K_{cd} \qquad (234) \\
& = \frac{(-\varepsilon)}{2\kappa c} \hat{G}^{abcd} K_{cd} \,.
\end{aligned}$$

Here again K stands for the expression (232). We also made use of the conformally rescaled De Witt metric (163) whose significance appears here for the first time. Again in passing we note that the physical dimension of π^{ab} is right, namely that of momentum per area (the dimension of K is an inverse length).

In order to compute the Hamiltonian density we express $\dot{h}$ in terms of the momenta

$$\begin{aligned}
\dot{h}_{ab} & = (L_\beta h)_{ab} - 2\varepsilon\alpha\, K_{ab} \\
& = (D_a\beta_b + D_b\beta_a) + 4\alpha\kappa c\, \hat{G}^{-1}_{abcd}\pi^{cd}
\end{aligned} \qquad (235)$$

and obtain

$$\begin{aligned}
\mathcal{H}_0[h,\pi] & = \pi^{ab} c\dot{h}_{ab} - \mathcal{L}_{\text{GR}} \\
& = \alpha\Big[(2\kappa c^2)\hat{G}^{-1}_{abcd}\pi^{ab}\pi^{cd} \\
& \quad + \epsilon(2\kappa)^{-1}\sqrt{h}R \Big] \\
& \quad + 2c\pi^{ab}D_a\beta_b \,.
\end{aligned} \qquad (236)$$

The Hamiltonian, H_0, is just the integral of this density over Σ. The subscript 0 is to indicate that this Hamiltonian is still to be modified by constraints according to the general scheme. Also, we have to once more care about surface terms in order to ensure functional differentiability without which the Hamiltonian flow does not exist [105].

The first thing to note is that we have found the primary constraints (233). For them to be maintained under the evolution we need to impose

$$c\dot{\pi}_\alpha = \{\pi_\alpha, H_0\} = -\frac{\delta H_o}{\delta\alpha} = 0, \qquad (237a)$$

$$c\dot{\pi}_{\beta^a} = \{\pi_{\beta^a}, H_0\} = -\frac{\delta H_o}{\delta\beta^a} = 0, \qquad (237b)$$

giving rise to the secondary constraints

$$(2\kappa c^2)\hat{G}^{-1}(\pi,\pi) + \varepsilon(2\kappa)^{-1}\sqrt{h}R = 0, \qquad (238a)$$

$$-2D_a\pi^{ab} = 0, \qquad (238b)$$

respectively. It may be checked directly that these equations respectively are equivalent to (144) for $\mathbf{T} = 0$. If we had included a cosmological constant this would have led to the replacement of R in (238a) with $(R - 2\Lambda)$.

In passing we make the following geometric observation regarding the bilinear form $G^{-1}(\pi,\pi)$, which is analogous to that made for $G(K,K)$ below equation (145). We can choose a local frame $\{e_a\}$ so that $h_{ab} = h(e_a, e_b) = \delta_{ab}$ and $\pi^{ab} = \pi(\theta^a, \theta^b) = \mathrm{diag}(p_1, p_2, p_3)$, where $\{\theta^a\}$ is dual to $\{e_a\}$. Then

$$\hat{G}^{-1}(\pi,\pi) := \left(\delta^{ab} - \tfrac{3}{2}n^a n^b\right)p_a p_b, \qquad (239)$$

where, as before, n^a are the components of the normalized vector $\vec{n} := (1,1,1)/\sqrt{3}$ in eigenvalue-space, which we identify with $\mathbb{R}^3$ endowed with the standard Euclidean inner product. Denoting again by θ the angle between $\vec{n}$ and $\vec{p} := (p_1, p_2, p_3)$, we have

$$\hat{G}^{-1}(\pi,\pi) = \begin{cases} 0 & \text{if} \quad |\cos\theta| = \sqrt{2/3} \\ > 0 & \text{if} \quad |\cos\theta| < \sqrt{2/3} \\ < 0 & \text{if} \quad |\cos\theta| > \sqrt{2/3}. \end{cases}$$
$$(240)$$

This should be compared with (147). The difference is that $\sqrt{1/3}$ is replaced by $\sqrt{2/3}$, which is due to $\lambda = 1$ in (157b) but $\mu = 1/2$ in (158b) in the GR case. This has an interesting consequence: The condition $|\cos\theta| = \sqrt{2/3}$ now describes a double cone around the symmetry axis generated by $\vec{n}$ and vertex at the origin, whose opening angle is strictly smaller than that of the cone considered

in (147). In fact, it is small enough to just touch boundaries of the positive and negative octants in $\mathbb{R}^3$. This means that $|\cos\theta| > \sqrt{2/3}$ implies that either all p_a are positive or all p_a are negative. In other words $\hat{G}^{-1}(\pi,\pi) < 0$ implies that the symmetric bilinear form π is either positive or negative definite. In contrast, (147) did not allow to conclude definiteness of the symmetric bilinear form K from $G(K,K) < 0$, since the interiors of the double-cone $|\cos\theta| > \sqrt{1/3}$ intersect the complements of the definite octants.

Let us now return to the constraints. We have found the primary and secondary constraints (233) and (238) respectively. The most important thing to note next is that there will be no further (tertiary, etc.) constraints. Indeed, this follows from the general argument following (150), which ensures the preservation of the secondary constraints under Hamiltonian evolution. The primary constraints are now taken care of by simply eliminating the canonical pairs (α, π_α) and (β^a, π_{β^a}) from the list of canonical variables. As we will see shortly, the secondary constraints (238) are of first-class, so that, according to the general theory outlined above, they should be added with arbitrary coefficients to the initial Hamiltonian H_0 to get the general Hamiltonian. This leads to

$$H[\alpha, \beta] = C_s(\alpha) + C_v(\beta) + \text{boundary terms}, \quad (241)$$

where

$$C_s(\alpha) := \int_\Sigma d^3x \; \alpha \; \Big[(2\kappa c^2)\hat{G}^{-1}(\pi,\pi)$$
$$+ \varepsilon(2\kappa)^{-1}\sqrt{h}R\Big], \quad (242a)$$

$$C_v(\beta) := \int_\Sigma d^3x \; \beta^a \Big[-2ch_{ab}D_c\pi^{bc}\Big], \qquad (242b)$$

where α and β^a are now arbitrary coefficients corresponding to the λ's in (180). In particular, they may depend on the remaining canonical variables h and π. Note that up to boundary terms the Hamiltonian is just a sum of constraints, where S stands for the scalar- (or Hamiltonian-) and and V for the vector (or diffeomorphism-) constraint.

The equations of motion generated by H will clearly be equivalent to (149) and (148). Let us

write them down explicitly. To do this we first note that the functional derivatives of $C_v(\beta)$ with respect to h and π are easily obtained if we note that, modulo surface terms, the integrand can either be written as $c(L_\beta h)_{ab}\pi^{ab}$ or as $-c(L_\beta \pi)^{ab}h_{ab}$. Hence, given that the surface terms are so chosen to guarantee functional differentiability, we have

$$\frac{\delta C_v(\beta)}{\delta \pi^{ab}} = c\,L_\beta h_{ab} = c\,(D_a\beta_b + D_b\beta_a) \qquad (243)$$

and

$$\frac{\delta C_v(\beta)}{\delta h_{ab}} = -c\,(L_\beta \pi)^{ab}$$
$$= -c\left[D_c(\beta^c\pi^{ab}) - (D_c\beta^a)\pi^{cb} - (D_c\beta^b)\pi^{ac}\right]. \qquad (244)$$

Note that in the first term on the right-hand side of (244) the β^c appears under the differentiation D_c because π is a tensor-density of weight one. The functional derivative with respect to π of $C_s(\alpha)$ is simply $4\alpha\kappa c^2\hat{G}(\pi, \cdot)$, so that the equation of motion for h is readily written down:

$$c\dot{h}_{ab} = \{h_{ab}, H\} = \frac{\delta H}{\delta \pi^{ab}} \qquad (245)$$
$$= 4\alpha\kappa c^2\hat{G}^{-1}_{ab\,cd}\pi^{cd} + cL_\beta h_{ab}\,.$$

Using (234) this is immediately seen to be just (149). From the explicit h-dependence of $\hat{G}$, as displayed in (158b) and (163), we obtain, using $\partial\sqrt{h}/\partial h_{ab} = \frac{1}{2}\sqrt{h}h^{ab}$,

$$\frac{\partial\hat{G}^{-1}(\pi,\pi)}{\partial h_{ab}} = -\frac{1}{2}h^{ab}\hat{G}^{-1}(\pi,\pi) + 2h^{an}\pi^{bm}\hat{G}^{-1}_{nmcd}\pi^{cd}. \qquad (246)$$

For the variational derivative of the second term in (242) we use the following standard formula for the variation of $\sqrt{h}$ times the scalar curvature of h

$$\delta\left(\sqrt{h}\,R(h)\right) = \sqrt{h}\left(-G^{ab}(h)\delta h_{ab} + G^{nmab}D_nD_m\delta h_{ab}\right), \qquad (247)$$

which immediately follows from (117), where we recall that g_{ab} there corresponds to h_{ab} here and also that $\delta h^{ab} = -h^{ac}h^{bd}\delta h_{cd}$. $G^{ab}(h)$ denote the

contravariant components of the Einstein tensor for h. Taken together we get

$$\frac{\delta C_s(\alpha)}{\delta h_{ab}} =$$
$$\kappa c^2\alpha\left[-h^{ab}\hat{G}^{-1}(\pi,\pi) + 4h^{an}\pi^{bm}\hat{G}^{-1}_{nmcd}\pi^{cd}\right]$$
$$-\frac{\varepsilon}{2\kappa}\left[\alpha G^{ab}(h) - G^{abnm}D_nD_m\alpha\right]\sqrt{h} \qquad (248)$$

With (244) this gives the second Hamilton equation

$$c\dot{\pi}^{ab} = \{\pi^{ab}, H\} = -\frac{\delta H}{\delta h_{ab}}$$
$$= \kappa c^2\alpha\left[h^{ab}\hat{G}^{-1}(\pi,\pi) - 4h^{an}\pi^{bm}\hat{G}^{-1}_{nmcd}\pi^{cd}\right]$$
$$+ \frac{\varepsilon}{2\kappa}\left[\alpha G^{ab}(h) - G^{abnm}D_nD_m\alpha\right]\sqrt{h}$$
$$+ cL_\beta\pi^{ab}\,. \qquad (249)$$

These are seen to be equivalent to (149) and (148) respectively. Before we discuss the boundary terms we write down the Poisson brackets for the constraints.

$$\{C_v(\beta), C_v(\beta')\} = C_v\left([\beta,\beta']\right), \qquad (250a)$$
$$\{C_v(\beta), C_s(\alpha)\} = C_s\left(\beta(\alpha)\right), \qquad (250b)$$
$$\{C_s(\alpha), C_s(\alpha')\} = \varepsilon C_v\left(\alpha(d\alpha')^\sharp - \alpha'(d\alpha)^\sharp\right). \qquad (250c)$$

These may be obtained by direct computation, but are also dictated by geometry. Before discussing the geometry behind them, we note the following more or less obvious points:

1. The vector constrains form a Lie algebra. The map $\beta \to V(\beta)$ is a Lie homomorphism form the Lie algebra of vector fields in Σ to the Lie algebra (within the Poisson algebra) of phase-space functions. In fact, this map is just the momentum map for the action of the diffeomorphism group $G = \text{Diff}(\Sigma)$ on phase T^*Q, which is a lift of the action on $Q = \text{Riem}(\Sigma)$, the space of Riemannian metrics on Σ. Note

39

that here the symplectic potential can be written in a symbolic infinite-dimensional notation (cf. (196a))

$$\theta = \int_{\Sigma} d^3x \; \pi^{ab}(x) \, \delta h_{ab}(x) \qquad (251)$$

and the vector field V_β generated by the action of $G = \text{Diff}(\Sigma)$ on $Q = \text{Riem}(\Sigma)$:

$$V_\beta = \int_{\Sigma} d^3x \; L_\beta h_{ab}(x) \frac{\delta}{\delta h_{ab(x)}} \,. \qquad (252)$$

The momentum map (222) is then given by

$$
\begin{aligned}
P_\beta = \theta(V_\beta) &= \int_{\Sigma} d^3x \; \pi^{ab} L_\beta h_{ab} \\
&= c^{-1} C_v(\beta) + 2 \int_{\partial\Sigma} d^2x \beta_a \pi^{ab} \nu_b \,,
\end{aligned}
\qquad (253)
$$

where ν^b denote the components of the outward pointing normal of $\partial\Sigma$. This shows that for vector fields β for which the surface term does not contribute the vector constraint is just the momentum map (up to a factor of c^{-1}, which comes in because the physical dimension of the values of the momentum map are that of momentum whereas the physical dimension of the constraints are that of an Hamiltonian, that is, energy). The surface term will be discussed below. What is important here is that the vector constraint coincides to the zero momentum-map for those diffeomorphisms which are asymptotically trivial, i.e. for which the surface term vanishes. Only those are to be considered as gauge transformations! Long ranging diffeomorphism for which the surface term is non zero, i.e. for configurations of non vanishing linear and/or angular momentum (cf. Section 9), have to be considered as proper changes in physical state. If we required these motions to be pure gauge we would eliminate all states with non-zero asymptotic charges. Compare the closing remarks of Section 7.

2. Once we have understood that the vector constraint is the momentum map for diffeomorphisms, its Poisson bracket with any other phase-space function F that defines a *geometric object* on Σ (i.e. an object with well defined transformation properties under diffeomorphisms) is fixed. We simply have

$$\{F, V_\beta\} = L_\beta F \,. \qquad (254)$$

In this sense (250b) says no more than that the expression (238a) is a scalar density of weight one. Recall that if F is a scalar density of weight one then $L_\beta F = D_a(\beta^a F)$. If we multiply F by α and integrate over Σ we get after partial integration and assuming the boundary term to give no contribution (which for non closed Σ requires certain fall-off conditions) an integral of $-F\beta(\alpha)$, which is just what (250b) expresses. Algebraically speaking, the fact that the Poisson bracket of a vector and a scalar constraint is proportional to a scalar rather than a vector constraint means that the vector constraints do *not* form an *ideal*. Geometrically this means that the Hamiltonian vector fields for the scalar constraint, if evaluated on the hypersurface for the vector constraint, will generally not be tangential to it, except for the points where this hypersurface intersects that of the scalar constraint. This has very important consequences for algorithms of phase-space reduction, i.e. algorithms that aim to "solve" the constraints. It means that a reduction in steps is *not* possible, whereby one *first* solves for the vector constraint and *then* seeks for solutions of the scalar constraint within the class of solutions to the vector constraint.

3. According to (250c) two scalar constraints Poisson commute into a vector constraint. Two facts are remarkable concerning the vector field that forms the argument of this vector constraint: First, it depends on the signature of spacetime (overall multiplication with ε). Second, it depends on the phase space variable h through the $\sharp$-operation of "index raising";

explicitly:

$$\alpha(d\alpha')^\sharp - \alpha'(d\alpha)^\sharp$$
$$= h^{ab}\big(\alpha\partial_b\alpha' - \alpha'\partial_b\alpha\big)\frac{\partial}{\partial x^a}\,. \qquad (255)$$

This is the fact already mentioned at the end of Section 7, that the constraints in GR are not altogether in the form of a vanishing momentum map. This fact has led to some discussion in the past and attempts have been made to consider different algebraic combinations of the constraints which define the same constraint hypersurfaces but display structure constants rather than structure functions in their Poisson brackets; e.g., [99]. But as already discussed in Section 7 it is important that these redefinitions do not spoil the regularity properties of the functions that define the constraint surface.

This ends the immediate discussion of (250). But there is another aspect that is related to the last point just discussed and that deserves to be mentioned.

8.1 Hypersurface deformations and their representations

Even though the constraints cannot be understood in a straightforward fashion as zero momentum map of a group action, they nevertheless do furnish a representation of an algebraic object (a groupoid) of hypersurface motions. As a result, the relations (250) are *universal*, in the sense that *any* spacetime diffeomorphism invariant theory, whatever its field content, will give rise to the very same relations (250); see [112] and [81] for early and lucid discussions and [88][89] for a comprehensive account.

The idea is to regard the space of (spacelike) embeddings $\mathrm{Emb}(\Sigma, M)$ of Σ into M as an infinite-dimensional manifold, on which the diffeomorphism group of M acts on the left by simple composition. Then there is a standard antihomomorphism from the Lie algebra of $\mathrm{Diff}(M)$ to the Lie algebra of vector fields on $\mathcal{E}(\Sigma, M)$, just as in (219). A tangent vector at a particular

$\mathcal{E} \in \mathrm{Emb}(\Sigma, M)$ can be visualized as a vector field ξ on $\Sigma \subset M$ with normal and tangential components, more precisely, as a section in the pull-back bundle $\mathcal{E}^* TM$ over Σ. Its decomposition into normal and tangential components depends on $\mathcal{E}$. If we think of M as being locally coordinatized by functions y^μ and Σ by functions x^a then $\mathcal{E}$ can be locally represented by four functions y^μ of three variables x^a. A vector field V_ξ can then be represented in a symbolic infinite-dimensional notation

$$V_\xi = \int_\Sigma d^3x\ \xi^\mu\big(y(x)\big)\frac{\delta}{\delta y^\mu(x)}\,. \qquad (256)$$

In full analogy to (219), this immediately leads to

$$[V_\xi, V_\eta] = -V_{[\xi,\eta]}\,. \qquad (257)$$

If we now decompose ξ in an embedding dependent fashion into its normal component αn and tangential component β we can rewrite (256) into

$$V(\alpha, \beta) =$$
$$\int_\Sigma d^3x\ \Big(\alpha(x)n^\mu[y](x) + \beta^a\partial_a y^\mu(x)\Big)\frac{\delta}{\delta y^\mu(x)}\,, \qquad (258)$$

where the components n^μ of the normal n to the image $\mathcal{E}(\Sigma) \subset M$ have to be considered as functional of the embedding. Again we can compute the commutator explicitly. The only non-trivial part is the functional derivative of the n^μ with respect to the y^ν. How this is done is explained in the Appendix of [112]. The result is

$$\big[V(\alpha_1, \beta_1), V(\alpha_2, \beta_2)\big] = -V(\alpha, \beta)\,, \qquad (259a)$$

where

$$\alpha = \beta_1(\alpha_2) - \beta_2(\alpha_1)\,, \qquad (259b)$$
$$\beta = [\beta_1, \beta_2] + \varepsilon\big(\alpha_1(d\alpha_2)^\sharp - \alpha_2(d\alpha_1)^\sharp\big)\,. \qquad (259c)$$

This is just (250) up to a relative minus sign that has the same origin as that between (219) and (223). We therefore see that (250) is a representation of a general algebraic structure which derives from the geometry of deformations of (spacelike) hypersurfaces in spacetime.

41

We can now address the inverse problem, namely to find all Hamiltonian representations of (259) on a given phase space. As in GR the phase space is T^*Q, where $Q = \mathrm{Riem}(\Sigma)$, That is, we may ask for the most general phase-space functions $H(\alpha, \beta)$: $T^*\mathrm{Riem}(\Sigma) \to \mathbb{R}$, parametrized by (α, β), so that

$$\{H(\alpha_1, \beta_1), H(\alpha_2, \beta_2)\} = H(\alpha, \beta)\,. \qquad (260)$$

The meaning of this relation is once more explained in Fig. 3. It is also sometimes expressed as *path independence*, for it implies that the Hamiltonian flow corresponding to two different paths in $\mathcal{E}(\Sigma, M)$ reaching the same final hypersurface will also result in the same physical state (phase-space point).
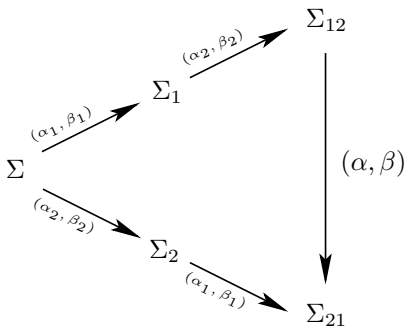
Figure 3: An (infinitesimal) hypersurface deformation with parameters (α_1, β_1) that maps $\Sigma \mapsto \Sigma_1$, followed by one with parameters (α_2, β_2) that maps $\Sigma_1 \mapsto \Sigma_{12}$ differs by one with parameters (α, β) given by (259b) from that in which the maps with the same parameters are composed in the opposite order.

To answer this question one first has to choose a phase space. Here we stick to the same phase space as in GR, that is T^*Q, where $Q = \mathrm{Riem}(\Sigma)$. The representation problem can be solved under certain additional hypotheses concerning the geometric interpretation of $H(\alpha = 0, \beta)$ and $H(\alpha, \beta = 0)$:

1. $H(0, \beta)$ should represent an infinitesimal spatial diffeomorphism, so that

$$\{F, H(0, \beta)\} = L_\beta F \qquad (261a)$$

for any phase-space function F. This fixes $H(0, \beta)$ to be the momentum map for the action of $\mathrm{Diff}(\Sigma)$ on phase space.

2. $H(\alpha, 0)$ should represent an infinitesimal $\mathrm{Diff}(M)$ action "normal to Σ". In absence of M, which is not yet constructed, this phrase is taken to mean that (80) must hold, i.e.

$$\{h, H(\alpha, 0)\} = -2\varepsilon\alpha K\,, \qquad (261b)$$

where K is the extrinsic curvature of Σ in the ambient spacetime that is yet to be constructed.

It has been shown that under these conditions the Hamiltonian of GR, including a cosmological constant, provides the unique 2-parameter family of solutions, the parameters being κ and Λ. See [81] for more details and [96] for the most complete proof (see below for a small topological gap). This result may be seen as Hamiltonian analog to Lovelock's uniqueness result [98] for Einstein's equations using spacetime covariance.

A particular consequence of this result is the impossibility to change the parameter λ in the De Witt metric (157) to any other than the GR value $\lambda = 1$ without violating the representation condition, that is, without violating covariance under spacetime diffeomorphisms. Such theories include those of Hořava-Lifshitz type [82], which were suggested as candidates for ultraviolet completions of GR.

At this point we must mention a topological subtlety which causes a small gap in the uniqueness proofs mentioned above and might have important consequences in Quantum Gravity. To approach this issue we recall from the symplectic framework that we can always perform a canonical transformation of the form

$$\pi \mapsto \pi' := \pi + \Theta\,, \qquad (262)$$

where Θ is a closed one-form on $\mathrm{Riem}(\Sigma)$. Closedness ensures that all Poisson brackets remain the same if π is replaced with π'. Since $\mathrm{Riem}(\Sigma)$ is an open positive convex cone in a vector space and hence contractible, it is immediate that $\Theta = df$

for some function $f : \mathrm{Riem}(\Sigma) \to \mathbb{R}$. However, π and π' must satisfy the diffeomorphism constraint, which is equivalent to saying that the kernel of π (considered as one-form on $\mathrm{Riem}(\Sigma)$) contains the vector fields generated by spatial diffeomorphisms, which implies that Θ, too, must annihilate all those, so that f is constant on each connected component of the $\mathrm{Diff}(\Sigma)$ orbit in $\mathrm{Riem}(\Sigma)$. But unless these orbits are connected this does not imply that f is the pull back of a function on the quotient $\mathrm{Riem}(\Sigma)/\mathrm{Diff}(\Sigma)$, as assumed in [96]. We can only conclude that Θ is the pull back of a closed but not necessarily exact one-form on superspace. Hence there is an analogue of the Bohm-Aharonov-like ambiguity that one always encounters if the configuration space is not simply connected. The quantum theory is then expected to display a sectorial structure labeled by the equivalence classes of unitary irreducible representations of the fundamental group of configuration space, which in analogy to Yang-Mills-type gauge theories are sometimes referred to as *θ-sectors* [87]. In GR the fundamental group of configuration space is isomorphic to a certain mapping-class group of the 3-manifold Σ. The theta-structure then depends on the topology of Σ and can range from 'trivial' to 'very complicated'. See [73] for more details on the role and determination of these mapping-class groups and [74] for a more general discussion of the configuration space in GR, which, roughly speaking, is the quotient $\mathrm{Riem}(\Sigma)/\mathrm{Diff}(\Sigma)$, often referred to as *Wheeler's superspace* [116][52].

We finally note that additional theta-structures may emerge if the gravitational field is formulated by means of different field variables including more mathematical degrees of freedom and more constraints (so as to result in the same number of physical degrees of freedom upon taking the quotient). The global structure of the additional gauge transformations may then add to the non-triviality of the fundamental group of configuration space and hence to the complexity of the sectorial structure. Examples have been discussed in the context of *Ashtekar variables* (cf. final Section) in connection with the *CP-problem* in Quantum Gravity [19].

8.2 An alternative action principle

A conceptually interesting albeit mathematically awkward alternative form of the action principle for GR was given by Baierlein, Sharp, and Wheeler in [21]. Its underlying idea, as far as the initial-value problem is concerned, is as follows: We have seen that initial data (h, K) (or (h, π)) had to obey four constraints (per space point) but that the four functions α and β could be specified freely. Could we not let the constraints determine α and β and thereby gain full freedom in specifying the initial data? In that case we would, for example, gain full control over the initial geometry, whereas, as we will see later, the standard *conformal method* to solve the constraints only provides control over the conformal equivalence class of the initial geometry, the representative within that class being determined by the solution to the scalar constraint. For black-hole collision data this, e.g., means that we cannot initially specify the initial distances. We will now discuss to what extent this can indeed be done. At the end of this subsection we will add some more comments regarding the conceptual issues associated with this alternative formulation.

We start from the action

$$S_{\mathrm{GR}}[g; \Omega] = \int_\Omega d^4 x \, \mathcal{L}_{\mathrm{GR}} \,, \qquad (263)$$

where $\mathcal{L}_{\mathrm{GR}}$ is as in (230). In it we express K in terms of $\dot{h}$ as given in (232). This results in

$$S_{\mathrm{GR}}[\alpha, \beta, h, \dot{h}\,; \Omega] =$$
$$\int_\Omega d^4 x \frac{\sqrt{h}}{2\kappa} \left\{ \frac{1}{4\alpha} \, G\big(\dot{h} - L_\beta h \,,\, \dot{h} - L_\beta h\big) - \varepsilon R \alpha \right\} . \tag{264}$$

For fixed domain $\Omega \subset M$ this is to be regarded as functional of g, that is of α, β and h. Note that α enters in an undifferentiated form. Variation with respect to it gives

$$\alpha = \alpha_*(h, \dot{h}, \beta) := \frac{1}{2} \sqrt{\frac{-\varepsilon \, G\big(\dot{h} - L_\beta h \,,\, \dot{h} - L_\beta h\big)}{R}} \,, \tag{265}$$

where we have chosen the positive root for α and introduced the abbreviation α_* for the function of h, $\dot{h}$, and β defined by the expression on the right-hand side in (265). Note that this only makes sense if R has no zeros and if the sign of $-\varepsilon\, G\big(\dot{h} - L_\beta h\,,\, \dot{h} - L_\beta h\big)$ equals that of R. Hence we have to restrict to a particular sign for the latter expression. We set

$$
\begin{aligned}
\sigma := &\ \mathrm{sign}\Big(G\big(\dot{h} - L_\beta h\,,\, \dot{h} - L_\beta h\big)\Big) \\
= &\ -\varepsilon\ \mathrm{sign}\big(R\big)\,.
\end{aligned}
\tag{266}
$$

Reinserting (265) into (263), taking into account (266), gives

$$
S_{\mathrm{BSW}}[\beta, h, \dot{h}; \Omega] := S_{\mathrm{GR}}[\alpha = \alpha_*\,,\, \beta, h, \dot{h}; \Omega] =
$$
$$
\int_\Omega d^4x\, \frac{\sigma\sqrt{h}}{2\kappa}\sqrt{-\varepsilon\, R\, G(\dot{h} - L_\beta h\,,\, \dot{h} - L_\beta h)}\,.
\tag{267}
$$

Here we explicitly indicated the functional dependence on the time derivative of h to stress the independence on the time derivative of β. With reference to [21] this form of the action is sometimes called the *Baierlein-Sharp-Wheeler action.* (Hence the subscript BSW in (267).) We can now try to further reduce this action so as to only depend on h and $\dot{h}$. For this we would have to proceed with β in the same fashion as we have just done with α; i.e., vary (267) with respect to β and then reinsert the solution $\beta[h, \dot{h}]$ of the ensuing variational equations back into (267). The variational equations for β are easily obtained (note that σ drops out):

$$
\begin{aligned}
&\frac{2\kappa}{\sqrt{h}}\, \frac{\delta S_{\mathrm{BSW}}}{\delta \beta_d} = \\
&\quad D_c\left\{ \frac{G^{abcd}(\dot{h} - L_\beta h)_{ab}}{\alpha_*(h, \dot{h}, \beta)} \right\} = 0\,.
\end{aligned}
\tag{268}
$$

These three equations are traditionally referred to as the *thin-sandwich equations.* They are meant to determine β for given pairs $(h, \dot{h})$.

Suppose there is a unique solution to (268) for given $(h, \dot{h})$, i.e.,

$$
\beta = \beta_*(h, \dot{h})\,.
\tag{269a}
$$

Inserting this into (265) determines the lapse for given $(h, \dot{h})$:

$$
\alpha = \alpha_*\big(h, \dot{h}, \beta = \beta_*(h, \dot{h})\big)\,.
\tag{269b}
$$

Our initial goal is then achieved if we consider $h \in \Gamma T_2^0 \Sigma$ (positive definite) and $\dot{h} \in \Gamma T_2^0 \Sigma$ (arbitrary) as freely specifiable initial data. As intended, h represents a freely specifiable Riemannian geometry of Σ and $\dot{h}$ its initial rate of change with respect to some formal parameter. The relation between this formal time parameter and proper time is fully determined by the solutions (269), in the way explained in Section 4.1. This means that the specification of two infinitesimally nearby configurations h and $h + \dot{h}\, dx^0$ allows to deduce the proper time that separates the corresponding spatial slices in the spacetime to be constructed. In this sense, and subject to the solvability of (268), physically meaningful *durations* can be deduced from two infinitesimally close instantaneous configurations. This is why Baierlein, Sharp, and Wheeler concluded in [21] that "three-dimensional geometry is a carrier of information about time"; or in Barbour's congenial dictum ([24], p. 2885): "The instant is not in time; time is in the instant". Also, it should now be obvious where the term *thin-sandwich* comes from and why the problem of finding a solution to (268) is often referred to as the *thin-sandwich problem.*

The thin-sandwich equations (268) were first conceived and discussed in [116] and [21]. Existence and uniqueness of solutions were originally conjectured (henceforth known as *thin-sandwich conjecture*), e.g. in [116], but first mathematical investigations soon showed that the unqualified thin-sandwich conjecture is false; see [30] and [41].

To see what positive statements can be made let us summarize the situation: At fixed time, i.e. on each 3-manifold Σ with given Riemannian metric h and given parameter-time derivative $\dot{h}$ (covariant symmetric second-rank tensor field), equations (268) form a system of three quasi-linear (though highly non-linear) second order equations for the three components β_a. The restriction $\alpha > 0$, where α is given by (265), involves β and hence implies an *a priori bound* for the unknown β. As a consequence one may first of all expect only *local* results

(if at all), in the sense that if β_* is a solution for given $(h_*, \dot{h}_*)$ then there exists some open neighborhood U of $(h_*, \dot{h}_*)$ in field space (in an appropriate topology) such that existence and uniqueness of solutions follows for all $(h, \dot{h}) \in U$. Such a local result can only be expected if the initial velocity $\dot{h}_*$, or any of its space-point dependent reparametrizations, is not just of the form of an infinitesimal diffeomorphism $L_\xi h_*$. In other words, $(h_*, \dot{h}_*)$ must be chosen such that for any smooth function η and vector field ξ on Σ we have the following implication

$$\dot{h}_* = \eta\, L_\xi h_* \quad \Leftrightarrow \quad \eta = 0 \quad \text{and} \quad \xi = 0\,. \quad (270)$$

Assuming (270), local results were first proven in [26] and subsequently in a more geometric form and with generalizations, also including matter fields, in [71]. The idea of proof is to write $\beta = \beta_* + \delta\beta$ and linearize the differential operator (268) acting on $\delta\beta$ around the solution $(\beta_*; h_*, \dot{h}_*)$. The resulting linear operator turns out to be symmetric (no surprise, being the second functional derivative of S_{BSW} in β) with a principal symbol $\sigma(k)$ whose determinant is proportional to a power of $\|k\|$ times $[\pi(k, k)]^2$ (see eq. (3.14) in [71]), which vanishes for non-zero k iff $\pi(k, k) = 0$. Hence the linearized operator is elliptic iff the quadratic form π is either positive or negative definite. Granted this, (270) then ensures that the elliptic operator has a trivial kernel. Together this allows to deduce an implicit-function theorem that immediately implies local existence and uniqueness. As regards ellipticity, we recall that from the discussion following (240) that a definite π is equivalent to $\hat{G}(\pi, \pi) < 0$, i.e. $\sigma = -1$ (compare (266)). Hence a local existence and uniqueness result holds provided that $(h_*, \dot{h}_*, \beta_*)$ satisfy $\sigma = -1$, comprising the two equations

$$\begin{aligned} &\epsilon R(h_*) > 0\,, \\ &G(\dot{h}_* - L_{\beta_*} h_*\,,\ \dot{h}_* - L_{\beta_*} h_*\,,) < 0\,. \end{aligned} \quad (271)$$

We see that we actually were not free in choosing either sign for σ in (266): We are bound to $\sigma = -1$ in order to ensure at least local existence and uniqueness of the thin-sandwich equation (268).

In passing we note the choice of the negative sign $\sigma = -1$ has very different topological consequences depending on whether $\varepsilon = -1$ (Lorentzian spacetime), where it implies $R < 0$, or $\varepsilon = 1$ (Riemannian/Euclidean spacetime), where it implies $R > 0$. In the former case, given any orientable closed 3-manifold Σ, the theorem of Kazdan & Warner (see paragraph above Section 6.1) ensures the existence of a Riemannian metric h on Σ with negative scalar curvature. In contrast, in the latter case (Euclidean spacetimes), the theorem of Gromov & Lawson implies a severe topological obstruction against Riemannian metrics h on Σ with positive scalar curvature (as already discussed above: only connected sums of handles and lens spaces survive this obstruction).

The result on local *uniqueness* can be generalized to a global argument, first given in [30] and generalized in [71] to also include matter fields. On the other hand, it is not difficult to see that global *existence* cannot hold, i.e. there are more or less obvious data $(h, \dot{h})$ for which (268) has no solution for β; see, e.g., [103].

Finally we mention the *conformal thin-sandwich* approach of York's [120], which is a conceptually weaker but mathematically less awkward variant of the full thin-sandwich problem, in which only the conformal equivalence class of the metric and its time derivative is initially specified, together with the lapse function and the extrinsic curvature. The constraints are then solved for the conformal factor and the shift vector field. The equation for the conformal factor is as in the conformal method discussed below (York's equation (328)), but the equation for the shift is mathematically less awkward than in the full thin-sandwich equation (268). Recall that the latter is rendered complicated to the insertion of the solution (265) for α, which is precisely what is *not* done in the conformal variant. But, clearly, the price for not solving for the lapse is that we have again no initial control over the full metric. Nevertheless, the better behaved equations of the conformal variant of the thin-sandwich method make it a useful tool in numerical investigations; see, e.g., [27][77].

Comparison with Jacobi's principle

It is conceptually interesting to compare the Baierlein-Sharp-Wheeler action of GR with Jacobi's principle in mechanics. So let us briefly recall Jacobi's original idea [90], where only the notation will be adapted.

As in Section 7 we consider a (so far unconstrained) mechanical system with n-dimensional configuration space Q. Let it be characterized by a Lagrangian $L : TQ \to \mathbb{R}$ of the form

$$L(q,v) = \tfrac{1}{2} G_q(v,v) - V(q) \,. \qquad (272)$$

Here G_q is a positive-definite bilinear form on $T_q Q$ called the *kinetic-energy metric*. We already know that as L does not explicitly depend on time, any dynamically possible trajectory will run entirely within a hypersurface of constant energy E (the energy function being (176)). Maupertuis' *principle of least action*, states that a dynamically possible trajectory $x : \mathbb{R} \supset I \to Q$, connecting fixed initial and final points q_i and q_f, extremizes the "action"

$$\int_{q_i}^{q_f} G_{x(t)}\big(\dot{x}(t), \dot{x}(t) \big) \, dt \qquad (273)$$

relative to all other curves with the same endpoints and on the same energy hypersurface. We note in passing that since $\dot{x} \, dt = dq$ and $G(v, \cdot) = \partial L(q,v)/\partial v = p$, the integrals in (273) equal the integrals of the canonical one-form θ (compare (196a)) along the paths $q = x(t)$ and $p = G\big(\dot{x}(t), \cdot \big)$ in T^*Q with fixed endpoints of the curve projected into Q and the curves all running on a hypersurface of fixed value of the Hamiltonian function. This is the form the principle of least action is given in modern formulations, like in Arnold's book [2](Chapter 9, Section 45 D). In this (modern, Hamiltonian) form, time plays no role. Indeed, on phase space T^*Q the integral of the one-form θ, as well as the level sets for the Hamiltonian function, are defined without reference to any time parameter t. But in the traditional (19th century, Lagrangian) form stated above, the parameter enters in an essential way. In fact, in this formulation t is not an *independent* variable because the energy

condition expressed in terms of positions and velocities (measured with respect to t) introduces an implicit dependency of t with coordinates on TQ. These dependencies have to be respected by variations of (273). This is why Jacobi complained in his lectures that the form given above appears incomprehensible to him.[1] His crucial observation was that the energy condition allows to eliminate t altogether. Indeed, if we solve the energy condition for dt,

$$dt = \sqrt{ \frac{G_q(dq,dq)}{2\big(E - V(q) \big)} } \,, \qquad (274)$$

and use that to eliminate dt in (273), we can put the integral into the form

$$\int_{q_i}^{q_f} \sqrt{ 2\big(E - V(q) \big) \, G_q(dq,dq) } \,, \qquad (275)$$

which is independent of any parametrization. In fact, it has the simple geometric interpretation of the length functional or the conformally rescaled kinetic-energy metric:

$$\hat{G} = 2(E - V) \, G \,, \qquad (276)$$

where E is a constant. *Jacobi's principle* then says that the dynamically possible trajectories of energy E are the geodesics of $\hat{G}$, and that (Newtonian) time along such a geodesic is obtained by integrating (274) along it. Note that Jacobi's principle defines in fact *two* new metrics on configuration space Q, both of which are conformally equivalent to the kinetic-energy metric G. The first is (276), which determines the trajectory in Q, the other is

$$\tilde{G} = \frac{G}{2(E - V)} \,, \qquad (277)$$

which determines Newtonian time along the trajectories selected by the first. We call it the *first* and *second Jacobi metric* respectively. Note also that (274) gives a measure for duration, dt, in terms of changes of mechanical coordinates.

---

[1] "Dieses Prinzip wird fast in allen Lehrbüchern, auch den besten, in denen von *Poisson, Lagrange* und *Laplace*, so dargestellt, dass es nach meiner Ansicht nicht zu verstehen ist." ([90], p. 44)

In passing we remark that a special realization of this far reaching idea, namely to read off Newtonian time from simultaneous configurations (i.e., generalized positions) of mechanical systems, assuming the systems to obey Newtonian laws of motion, is *emphemeris (or astronomical) time* in astronomy [47][23], where the relative configurations (ephemerides) of the Moon, Sun, and planets (as seen from the Earth) are used as positions. Ephemeris time was first proposed as time standard in 1948 [46] in order to establish a reference with respect to which non uniformities in the Earth's daily rotation could be accounted for, though the idea goes back at least to 1929 [48]. Ephemeris time ideed became the time standard in 1952 until the 1970s, when atomic time took over its place, though we remark that "atomic time" is really just based on a straightforward generalization of the very same principle to Quantum Mechanics. Here, again, one reads off time from simultaneously measurable states or observables of one or more systems obeying known deterministic "laws of motion", like Schrödinger's equation for states or Heisenberg's equation for observables. These equations of motion correlate the a priori unobservable parameter t with observable properties of the system, which render t observable in a context dependent fashion through inverting these relations (solving the equations of motion for t).

Now, coming back to gravity, (275) should be compared to (267). Except for the terms involving the lapse-function β the latter is (almost!) like (275) for $E = 0$ and $V = \varepsilon R$. The analogy is incomplete (hence "almost") because in (267) the spatial integration is outside the square-root, so that the integrand for the parameter-integration along the curve in configuration space is a sum of square-roots rather than a single square-root of a sum. This difference renders the expression in (267) different from usual "length functionals". Note that sums of square roots (involving themselves sums of squares) generally do not even form a Finsler metric (compare the discussion in [71]).

Taking the analogy further, (274) should be compared to (269b) with α_* given by the integral (265). Again, except for the terms involving β, they seem

to closely correspond to each other. Equation (269b) determines one proper time per spatially fixed (with respect to the spatial coordinates) observer in the spacetime to be developed from the initial data. Hence there is something like a continuum of second Jacobi metrics, one for each space point.

An interesting observation in this connection is the following [40], which we give in a simplified form. Suppose that we tried to define, from first principles, *duration* by some measure of change in the gravitational degrees of freedom, i.e., some kind of gravitational *ephemeris time* . In analogy to (274) We assume the "measure of change", $d\tau$, to be given by some local rescaling of a pseudo-Riemannian distance measure

$$ ds^2 = \int_\Sigma d^3x \, G^{ab\,cd}[h(x)]dh_{ab}(x)dh_{cd}(x) \,, \quad (278) $$

so that

$$ d\tau^2 = \frac{ds^2}{\int_\Sigma d^3x \, R(x)} \,. \quad (279) $$

Here R must be a scalar function of the spatial metric h. The simplest non-constant such function is the scalar curvature, which depends on h and its derivatives up to second order. A priori such a measure of duration seems to depend equally on all gravitational degrees of freedom at all points in space, thus giving rise to a highly non-local concept of time with respect to which durations of processes, even local ones, can be measured. However, suppose we required that the measure of time be compatible with arbitrarily fine localization $\Sigma \to U \subset \Sigma$. Following [40] we call this the *chronos principle*. It implies that the numerator and denominator of (279) are proportional for each restriction $\Sigma \to U$. This is only possible if the integrands are proportional. Without loss of generality we can take this constant of proportionality (which cannot be zero) to be 1 (this just fixes the overall scale of physical time) and obtain (here written without cosmological constant for simplicity)

$$ G^{ab\,cd}[h(x)]\frac{dh_{ab}(x)}{d\tau}\frac{dh_{cd}(x)}{d\tau} - R[h](x) = 0 \,. \quad (280) $$

which is just the Hamiltonian constraint. The duration in time is then given by

$$\Delta\tau(h_i, h_f) = \int_{h_i}^{h_f} \sqrt{\frac{G(dh/d\lambda, dh/d\lambda)}{-R[h(\lambda)]}}\, d\lambda\,, \quad (281)$$

which is obviously just the analog of the integrated version of (274) for $E = 0$ (which here amounts to the Hamiltonian constraint), i.e. the integrated version of (265) for $\beta = 0$ (this is the simplification we alluded to above) and $\varepsilon = -1$.

9 Asymptotic flatness and global charges

Isolated systems are described by geometries which at large spatial distances approach a matter-free spacetime. In case of vanishing cosmological constant the latter will be flat Minkowski spacetime. For non-zero Λ it will either be de Sitter ($\Lambda > 0$) or anti-de Sitter ($\Lambda < 0$) space. Here we are interested in the case $\Lambda = 0$. We refer to the survey [60] for a recent discussion of the anti-de Sitter case.

An initial data set (h, π) or (h, K) on Σ needs to satisfy certain asymptotic conditions in order to give rise to an asymptotically flat spacetime. Before going into this, we point out that there is also a topological condition on Σ in order to sensibly talk about *asymptotic regions*. The condition is that there exists a compact set $K \subset \Sigma$ such that its complement $\Sigma - K$ is diffeomorphic to the disjoint union of manifolds $\mathbb{R}^3 - B$, where B is a closed ball. These pieces in which Σ decomposes if one cuts out increasingly large compact sets are called *ends* of Σ. In passing we note that the theory of 'ends' for topological spaces and groups was developed by Freudenthal in 1931 [61]. Now, the first condition we pose is that there is only a finite number of such ends. (It is not hard to visualize manifolds with even an uncountable number of ends.) With respect to each end we can talk of approaching infinity. This means to let $r \to \infty$ if r is the standard radial coordinate on $\mathbb{R}^3 - B$ to which this end is diffeomorphic.

A first working definition of asymptotically flat initial data in Hamiltonian GR was given in 1974 by Regge & Teiltelboim [105]. It was shown by Beig & Ó Murchadha in 1986 [29] that this definition is sufficient to allow the implementation of the 10-parameter Poincaré group as asymptotic symmetries giving rise to 10 corresponding conserved quantities. The definition can be given as follows (here we restrict to one end):

Definition (Regge-Teitelboim asymptotic flatness). *Let Σ be a 3-manifold with one end. An initial data set (h, π) on Σ is asymptotically flat in the sense of Regge-Teitelboim if there is a coordinate system $\{x^1, x^2, x^3\}$ covering the end, such that as $r := \sqrt{x^a x^b \delta_{ab}} \to \infty$)*

$$h_{ab}(x) = \delta_{ab} + \frac{s_{ab}(\nu)}{r} + O_2(r^{-2})\,, \quad (282a)$$

$$\pi^{ab}(x) = \frac{t^{ab}(\nu)}{r^2} + O_1(r^{-3})\,, \quad (282b)$$

where $x = (x^1, x^2, x^3)$ and $\nu = (\nu^1, \nu^2, \nu^3)$ with $\nu^a := x^a/r$. $O_k(r^{-n})$ denotes terms falling off like $1/r^n$ and whose l-th derivatives fall off like $1/r^{(n+l)}$ for $0 < l \leq k$. Moreover, s_{ab} and t^{ab} obey the parity conditions

$$s_{ab}(-\nu) = s_{ab}(\nu)\,, \quad (283a)$$

$$t^{ab}(-\nu) = -t^{ab}(\nu)\,. \quad (283b)$$

The first thing to observe is that these conditions suffice to make the integral (251) for the symplectic potential convergent. Note that (282) merely implies that the integrand falls off like $1/r^3$, which could still produce a logarithmic divergence. But (283) implies that the $1/r^3$ integrand is of odd parity and hence gives no contribution. Next we have a look at the constraint functionals (242). In (242a) the first integrand has a $1/r^4$ and the second a $1/r^3$ parity-even fall-off. In (242b) the integrand has also a $1/r^3$ parity-even fall-off. Hence the integrals (242) certainly converge for those lapse and shift fields α, β which asymptotically either tend to zero or approach direction-dependent constants in a parity-odd fashion. As we will see below, the constraints for such parameter fields α and β are

differentiable with respect to the canonical variables h and π and hence generate a Hamiltonian flow that has to be considered as gauge transformations; compare Section 7.1. Hence we set

$$\alpha(x)_{\text{gauge}} = a(\nu) + O_2(r^{-1}), \qquad (284\text{a})$$

$$\beta^a(x)_{\text{gauge}} = b^a(\nu) + O_1(r^{-1}), \qquad (284\text{b})$$

where

$$a(-\nu) = -a(\nu), \qquad (285\text{a})$$

$$b^a(-\nu) = -b^a(\nu). \qquad (285\text{b})$$

To see that $C_s(\alpha)[h, \pi]$ and $C_v(\beta)[h, \pi]$ as defined in (242a) and (242b) are functionally differentiable with respect to h and π we make the following observations: For (242a) the only boundary term one might pick up is that from the variation of the scalar curvature with respect to h, which follows from (117) (in which formula we have to replace g_{ab} with h_{ab} and h_{ab} with δh_{ab} in order to match the notation here). It reads

$$\frac{\varepsilon}{2\kappa} \int_{S_\infty^2} d\Omega \, \alpha \sqrt{h} \nu_a G^{abcd} D_b \delta h_{cd} \qquad (286)$$

and is thus seen to have an integrand that has $1/r^2$ fall-off and is parity-odd. Hence the integral vanishes. Note that here and in what follows we used the following shorthand notation

$$\int_{S_\infty^2} d\Omega(\cdots) := \lim_{r \to \infty} \left\{ \int_{S^2(r)} d\Omega_r(\cdots) \right\}, \quad (287)$$

where $S^2(r)$ is the sphere of constant "radius" r (as defined above) and $d\Omega_r$ its induced volume form.

The vector constraint (242b) contains π as well as h in differentiated form (the latter in D), so that boundary terms may appear in both variation, that with respect to π as well as that with respect to h. For both cases it is convenient to rewrite the integral (242b) by performing a partial integration

before variation:

$$C_v(\beta) := \int_\Sigma d^3x \, \beta^a \Big[-2ch_{ab} D_c \pi^{bc} \Big]$$

$$= c \int_\Sigma d^3x (L_\beta h)_{ab} \pi^{ab} \qquad (288)$$

$$- 2c \int_{S_\infty^2} d\Omega \, \beta^a h_{ab} h_{cd} \nu^d \pi^{bc} .$$

Under the fall-offs and parity conditions mentioned above the last (surface) integral is zero since its integrand has $1/r^2$ decay and is parity-odd. Hence variation with respect to π does not lead to surface terms. Variation with respect to h is now simply given by varying the h under the Lie differentiation L_β which itself has no dependency on h (unlike the covariant derivative D). Using $L_\beta(\delta h_{ab}\pi^{ab}) = D_c(\beta^c \delta h_{ab}\pi^{ab})$, partial differentiation with respect to the Lie derivative then gives the surface term

$$2c \int_{S_\infty^2} d\Omega \, h_{ab}\nu^a \beta^b \, \delta h_{cd}\pi^{cd}, \qquad (289)$$

the integrand of which again falls off like $1/r^2$ and is parity odd.

The considerations so far show that the constraints $C_s(\alpha)[h, \pi]$ and $C_v(\beta)[h, \pi]$ are differentiable with respect to h and π whenever (h, π) satisfy (282) and (283) and the parameter-functions (α, β) satisfy (284) and (285). From the considerations it also follows that we cannot improve on the latter two conditions given the fall-offs and parity conditions on (h, π). This characterizes the lapse and shift functions which generate pure *gauge transformations* in Hamiltonian gravity for asymptotically states. We stress that (284) and (285) includes motions that do not vanish at infinity. These are called *supertranslations*. Without their careful inclusion into the transformations considered as gauge, we would not obtain the Poincaré group as proper physical symmetry group but rather an infinite-dimensional extension thereof. For more discussion on this conceptually important point compare the discussion in [69].

Motions characterized by functions (α, β) outside the class (284) and (285) do change the physical state. If this motion is to be generated by

the Hamiltonian (241) we must restrict to those (α, β) for which suitable boundary terms can be found such that $H(\alpha, \beta)[h, \pi]$ is differentiable with respect to h and π obeying (282) and (283). To accommodate asymptotic Poincaré transformations we must worry about asymptotic translations in time and space directions, asymptotic rotations, and finally asymptotic boosts, all of which we only need to specify modulo gauge transformations. Asymptotic time translations correspond to constant α. The surface term that results from the variation of $C_s(\alpha)$ is just (286). It immediately follows that the term that has to be added to $C_s(\alpha)$ so as to just cancel this surface term upon variation with respect to h is just $\alpha\, E_{\mathrm{ADM}}$, where

$$E_{\mathrm{ADM}} = M_{\mathrm{ADM}} c^2$$
$$= -\varepsilon (2\kappa)^{-1} \int_{S^2_\infty} d\Omega \, (\partial_a h_{ab} - \partial_b h_{aa}) \nu_b \,,$$
$$(290)$$

is called the *ADM energy* and M_{ADM} the *ADM mass*. Note that we just replaced all non differentiated h_{ab} that appear in (286) by δ_{ab} since the difference does not contribute to the surface integral in the limit as $r \to \infty$. Similarly, asymptotic space translations corresponds to covariantly constant β, i.e. constant components β^a with respect to the preferred coordinates that served to define asymptotic flatness. Again we immediately read off (288) the boundary term that we need to add in order to cancel that in the last line of (288) upon variation of π. It can be written in the form $c P_{\mathrm{ADM}}(\beta)$, where

$$P_{\mathrm{ADM}}(\beta) = 2 \int_{S^2_\infty} d\Omega \, \beta^a \pi_{ab} \nu^b \,. \qquad (291)$$

This we call the *ADM general momentum*. Note that the integrand has fall-off $1/r^2$ and even parity and hence gives a finite contribution. Furthermore, it follows from (289) that the variation with respect to h does not give rise to a boundary term since the integrand in (289) has $1/r^3$ fall off and hence does not contribute, independently of its (even) parity.

Asymptotic rotations and boosts are a priori more delicate since now α and β are allowed to grow linearly with r. We have, up to gauge transformations, $\alpha \propto u_a x^a$ for a boost in $\vec{u}$-direction and $\beta_a \propto \varepsilon_{abc} \omega^b x^c$ for a rotation around the $\vec{\omega}$ axis. Here ε_{abc} are the components of the metric volume-form for the asymptotic metric δ with respect to the coordinates $\{x^a\}$, so that $\varepsilon_{abc} = \pm 1$, depending on whether (abc) is an even $(+)$ or odd $(-)$ permutation of (123). As indices are raised and lowered with respect to δ, we need not be concerned whether they are upper or lower (as long as we work in components with respect to $\{x^a\}$). The components u_a and ω^a are then again constants.

We start with rotations and read off the last line of (288) that the surface integral has an integrand $\propto \pi^{ab} \beta_a \nu_b$, which looks dangerous as its naive fall-off is $1/r$. However, if we use that π actually satisfies the constraint, $D_b \pi^{ab} = 0$, we can convert this surface integral to a bulk integral whose integrand is proportional to $D_b(\pi^{ab} \beta_a) = \pi^{ab} D_{(a} \beta_{b)}$. But $D_{(a} \beta_{b)} = \partial_{(a} \beta_{b)} - \Gamma^c_{ab} \beta_c$ and β is Killing with respect to the metric δ_{ab}, so that $\partial_{(a} \beta_{b)} = 0$. Now, the Christoffel symbols Γ^c_{ab} for the metric h fall off as $1/r^2$ with odd parity, so that $\Gamma^c_{ab} \beta_c$ falls off as $1/r$ with even parity. Hence $\pi^{ab} D_{(a} \beta_{b)}$ falls off like $1/r^3$ with odd parity, showing that this volume integral also converges (a logarithmic divergence being just avoided by odd parity). Finally we observe that for asymptotic rotations there is still no surface term of the form (289), since its integrand has $1/r^2$ fall-off and is of odd parity. As a result we have that even for asymptotic rotations we obtain the same formula (291) for the (linear or angular) momentum as long as the Regge-Teitelboim conditions (282) and (283) are satisfied and (h, π) satisfy the constraints. In components with respect to the asymptotic coordinates the components of the linear momentum are

$$P^a_{\mathrm{ADM}} = 2 \int_{S^2_\infty} d\Omega \, \pi^{ab} \nu_b \,, \qquad (292)$$

and for the angular momentum

$$J^a_{\mathrm{ADM}} = 2 \int_{S^2_\infty} d\Omega \, \varepsilon_{abc} x^b \pi^{cd} \nu_d \,, \qquad (293)$$

where ε_{abc} is as above.

50

Turning now to the boosts, we have

$$\alpha = u_a x^a + \alpha_{\text{gauge}} \, . \tag{294}$$

We need to repeat the same steps that previously led us to (286). Now, according to (117) the space integral over the divergence term in the variation of the scalar curvature is

$$X := \int_\Sigma d^3x \, \alpha \sqrt{h} G^{abcd} D_a D_b \delta h_{cd} \, . \tag{295}$$

A first integration by parts leads to

$$X = \int_{S^2_\infty} d\Omega \, \alpha \sqrt{h} \nu_a G^{abcd} D_b \delta h_{cd} \\ - \int_\Sigma d^3x \, \sqrt{h} G^{abcd} (D_a \alpha) D_b \delta h_{cd} \, . \tag{296}$$

One more integration by parts of the second term gives

$$X = \int_{S^2_\infty} d\Omega \, \alpha \sqrt{h} \nu_a G^{abcd} D_b \delta h_{cd} \\ - \int_{S^2_\infty} d\Omega \, \sqrt{h} \, (D_a \alpha) \nu_b G^{abcd} \delta h_{cd} \\ + \int_\Sigma d^3x \, \sqrt{h} G^{abcd} (D_a D_b \alpha) \delta h_{cd} \, . \tag{297}$$

Equation (294) implies that $D^2\alpha$ has $1/r^2$ fall-off with odd parity. Hence the last (volume) integral in (297) has $1/r^3$ fall-off with odd parity and hence converges. It gives rise to a term $\propto \sqrt{h} G^{abcd} D_a D_b \alpha$ in the Hamiltonian equation for $\dot{\pi}^{cd}$. According to the general strategy the surface integrals must be taken care of by adding suitable surface integrals to the right-hand side of (241) so as to just cancel $\varepsilon/(2\kappa)$ times the integrals just found as resulting from the variation of the scalar curvature in (242). Hence the right surface terms to be added to (241) are of the form $u_a X^a$, where

$$X^a = \frac{-\varepsilon}{2\kappa} \left\{ \int_{S^2_\infty} d\Omega \, x^a \big(\partial_b h_{bc} - \partial_c h_{bb} \big) \nu_c \\ - \int_{S^2_\infty} d\Omega \, \big((h_{ab} - \delta_{ab}) \nu_b - (h_{bb} - \delta_{bb}) \nu_a \big) \right\} \, . \tag{298}$$

The coordinates Z^a of the *center of mass* are then defined by the rescaled forms of (298), with rescaling factor E_{ADM}:

$$Z^a := \frac{X^a}{E_{\text{ADM}}} \, . \tag{299}$$

In order to arrive at (298) we wrote $\delta h_{ab} = \delta(h_{ab} - \delta_{ab})$ and left the difference $(h_{ab} - \delta_{ab})$ rather than just h_{ab} under the integral in order to not keep the asymptotically constant term of (282a) under the integral when pulling the variation δ outside it (cf. [29] Appendix C).

It has been shown in [49] that the expression (298) for the (unscaled) center of mass coincides with the geometric definition of Huisken and Yau's [84]. The latter is defined by means of mean-curvature foliations of Σ. Its relation to alternative definitions, including not only ADM but also a definition due to R. Schoen, using asymptotically conformal Killing fields, is discussed and lucidly summarized in [83].

So far we have been working with the particular asymptotic conditions (282) and (283). We have been arguing for existence of certain quantities to be identified with physical quantities of energy, linear and angular momentum, and center of mass. But what about uniqueness? All these quantities depend a priori on the choice of the asymptotic coordinates within the set of all coordinates satisfying the given fall-off conditions. Hence one needs to prove that this dependence is actually spurious and that, consequently, these quantities are geometric invariants. For the ADM mass and linear momentum this has been shown in [25] and [43]. Moreover, ignoring angular momentum and center of mass, these proofs were given under much weaker asymptotic conditions, in fact the weakest possible ones. Regarding the latter, we recall that it was shown in [51] by means of explicit coordinate transformations that the expression can be made change its value if $|h_{ab} - \delta_{ab}| < r^{-\alpha}$ with $\alpha \le 1/2$; see also the lucid discussion in [50]. Hence we certainly need $\alpha > 1/2$. That this indeed suffices to prove existence and uniqueness was established in [102, 25, 43]. This fits nicely with recent generalizations of stability results of Minkowski space by

Bieri [31], which work under the following asymtotic decay conditions, where $\alpha > 1/2$:

$$h_{ab} = \delta_{ab} + O_2(r^{-\alpha}), \tag{300a}$$

$$\pi^{ab} = O_1(r^{-1-\alpha}). \tag{300b}$$

These conditions suffice to establish ADM energy and linear momentum not only as being well defined, but also as being preserved under Hamiltonian evolution.

At first sight (300a) might seem too weak to guarantee existence of (290). The reason why it is not is, in fact, easy to see: If we convert (290) into a bulk integral using Gauss' theorem, the integrand contains a combination of 2nd derivatives of h which just form the 2nd derivative part of the scalar curvature. Using the scalar constraint, which schematically has the form

$$\pi^2 + \partial^2 h + (\partial h)^2 = 0, \tag{301}$$

this can be written as a bulk integral containing only integrands of the form $\propto \pi^2$ and $\propto (\partial h)^2$, which according to (300a) fall off like $1/r^{2(1+\alpha)}$, i.e., faster than $1/r^3$. Hence the bulk integral converges. But note that the conditions (300) do not suffice to ensure the existence of conserved quantities regarding angular momentum or center of mass. Alternative conditions to (282) and (283) for the existence of angular momentum have been discussed in [20] and [44].

We recall that even in the context of the Regge-Teitelboim conditions (282) and (283) we needed to invoke the fact that (h, π) satisfy the constraints in order to conclude sufficiently strong fall-offs. This we have already seen explicitly in the discussion on, e.g., the existence of angular momentum in the paragraph above equation (292). From the scalar constraint (301) we now learn that (282) and (283) implies a $1/r^4$ fall-off for $G^{abcd}\partial_a\partial_b h_{cd}$, and not just $1/r^3$ as naively anticipated from (282).

Alternative expressions for mass/energy exist in cases of symmetries. For example, for asymptotically flat and *stationary* solutions to Einstein's equations, the ADM mass M_{ADM} is known to coincide with the so-called Komar mass [94], whose simple and coordinate invariant expression is

$$M_{\mathrm{Komar}} = \frac{-\varepsilon}{c^2\kappa} \int_{S^2_\infty} \star dK^\flat. \tag{302}$$

Here K is the timelike Killing vector field so normalized that $\lim_{r\to\infty} g(K,K) = \varepsilon$. There exist various proofs in the literature showing $M_{\mathrm{Komar}} = M_{\mathrm{ADM}}$; see, e.g., [28][20][43] and Theorem 4.13 of [39].

Since the Komar mass is frequently used in applications, let us say a few things about it. From a mathematical point of view the main merit of (302) is that it allows to associate a "mass" to any 2-dimensional submanifold $S \in M$, independently of any choice of coordinates. We call it the Komar mass of S:

$$M_{\mathrm{Komar}}(S) = \frac{-\varepsilon}{c^2\kappa} \int_S \star dK^\flat. \tag{303}$$

For $S \to S^2_\infty$ its interpretation is that of the ADM mass, whose physical significance as the value of the Hamiltonian (divided by c^2) endows it with a sound physical interpretation. But what might the interpretation be for general S? Well, suppose $S = \partial B$, where B is a 3-dimensional spacelike submanifold of M. Then, by Stokes' theorem:

$$\begin{aligned} M_{\mathrm{Komar}}(S) &= \frac{-\varepsilon}{c^2\kappa} \int_B d \star dK^\flat \\ &= \frac{1}{c^2\kappa} \int_B \star(\star d \star dK^\flat) \\ &= \frac{\varepsilon}{c^2\kappa} \int_B \star(\nabla \cdot dK^\flat) \\ &= \frac{-2\varepsilon}{c^2\kappa} \int_B \star i_k \mathbf{Ric}. \end{aligned} \tag{304}$$

Here we used the general identity for the square of the Hodge star restricted to p-forms in n dimensions,

$$\star \circ \star \big|_{\Lambda^p(M)} = \varepsilon\,(-1)^{p(n-p)}\,\mathrm{Id}_{\Lambda^p(M)}, \tag{305}$$

and also the general formula that allows to express $\star d \star$ in terms of the covariant divergence $\nabla \cdot$ on the

first index with respect to the Levi-Civita connection,

$$\star \circ d \circ \star\big|_{\Lambda^p(M)} = \varepsilon \, (-1)^{n(p+1)} \, \nabla \cdot . \qquad (306)$$

In the final step of (304) we used that any Killing vector-field satisfies the identity (again for the Levi-Civita connection):

$$\nabla_a \nabla_b K_c = R^d_{\ abc} K_d . \qquad (307)$$

If the spacetime satisfies Einstein's equation we can use (5) to eliminate the Ricci tensor in the last line of (303) in favor of the energy-momentum tensor. This shows that if S has two connected components S_1 and S_2, and if $\mathbf{T}|_B = 0$, then (choosing the relative orientations of S_1 and S_2 appropriately) $M_{\text{Komar}}(S_1) = M_{\text{Komar}}(S_2)M$. For a finite-size star or a black hole this means that we may take any S to calculate the Komar- and hence the ADM mass, as long as $S \cup S_\infty^2$ bounds a 3-dimensional region B on which $\mathbf{T}$ vanishes. In particular, S may be taken as any 2-sphere outside the star's surface. More specifically, consider a static star where K is the hypersurface orthogonal Killing vector-field. The topology of the hypersurfaces orthogonal to K inside the star shall be just that of a ball in $\mathbb{R}^3$; then we express the star's ADM energy by the Komar integral over the star's surface S and that, in turn, by the bulk integral (304) over the star's interior, where we replace the Ricci tensor by $\mathbf{T}$. This results in the so-called Tolman mass (see [114] and §92 of [115]), which in our notation reads:

$$M_{\text{Tolman}} = \frac{-\varepsilon}{c^2} \int_b d^3x \ \sqrt{\det(h)} \times$$
$$\sqrt{\varepsilon g(K,K)} \left[\mathbf{T}(n,n) - \varepsilon \text{Tr}_h(\mathbf{T}) \right]. \qquad (308)$$

Here $n := K/\sqrt{\epsilon g(K,K)}$ is the normal to the hypersurfaces and Tr_h the trace with respect to the spatial metric h, where, we recall, $g = \varepsilon \, n^\flat \otimes n^\flat + h$. In the Lorentzian case ($\varepsilon = -1$) we see that the first $T(n,n)$-term in (308) is just the integral over the spatial energy-density of the matter divided by c^2 and weighted by the redshift factor $\sqrt{-g(K,K)}$. The additional term is absent if the

pressures are negligible compared to the energy density, but this need not be the case. For example, if $\mathbf{T}$ is that of an electromagnetic field, we have $\text{Tr}_g(\mathbf{T}) = \varepsilon T(n,n) + \text{Tr}_h(\mathbf{T}) = 0$, so that the pressure effectively doubles the contribution of the first term to the overall mass. This is the origin of the infamous "factor-2-anomaly" of the Komar mass, which, e.g., leads to the result that the difference between two Komar masses evaluated on two different 2-spheres of spherical symmetry in the Reissner-Nordstrøm manifold (electrically charged black-hole) gives *twice* the electrostatic field energy stored in the region between the spheres. On the other hand, for a spherically symmetric perfect-fluid star, Tolman has shown in [114] that in a weak-field approximation the leading-order difference between (308) and the integrated mass-density of the fluid is just the Newtonian binding energy, which makes perfect sense.

At this point we should mention the positive-mass theorem (for Lorentzian signature $\varepsilon = -1$), which states that for any pair (h, π) of initial data satisfying the constraints $M_{\text{ADM}} \geq 0$, with equality only if the data are that of Minkowski space. Note that the expression (290) for M_{ADM} is a functional of h alone, but that in the formulation of the positive-mass theorem given here it is crucial that for h there exists a π so that the pair (h, π) solves the constraint. Otherwise it is easy to write down 3-metrics with negative ADM mass; take e.g. (314) (see below) with r_0 replaced by $-r_0$, where $r_0 > 0$, suitably smoothed out for smaller radii so as to avoid the singularity at $r = r_0$. Since the ADM mass only depends on the asymptotic behavior it is completely independent of any alterations to the metric in the interior. If one wishes to make the positive mass theorem a statement about metrics alone without any reference to the constraints one has to impose positivity conditions on the scalar curvature. But that also imposes topological restrictions due to the result of Gromov and Lawson [78] mentioned at the end of Section 6. For a recent up-to-date survey on the positive-mass theorem we refer to [50].

Note that $M_{\text{ADM}} = M_{\text{Komar}}$ implies the positivity of $M_{\text{Komar}} \equiv M_{\text{Komar}}(S_\infty^2)$. But this does

not imply that $M_{\mathrm{Komar}}(S^2)$ is also positive for general S^2. In fact, explicit examples of regular, asymptotically flat spacetimes with matter satisfying the hypotheses of the positive-mass theorem are known in which $M_{\mathrm{Komar}}(S^2) < 0$ for suitably chosen 2-spheres [1]. The recipe here is to regard two concentric counter-rotating objects in an axially-symmetric and stationary spacetime, e.g., an outside perfect-fluid ring and an inner rigidly rotating disk of dust. The Komar mass of the disk (i.e. S^2 encloses the disk but not the ring) may then turn out negative if the frame-dragging effect of the ring is large enough so as to let the angular velocity and the (Komar) angular momentum of the central object have *opposite* signs.

In passing we remark that the positive mass theorem in combination with the equality $M_{\mathrm{ADM}} = M_{\mathrm{Komar}}$ gives a simple proof of the absence of "gravitational solitons", i.e. stationary asymptotically flat solutions to Einstein's equations on $\Sigma = \mathbb{R}^3$. This follows from (302) and $d \star dK^\flat \propto \star i_K \mathbf{Ric}$. The vaccum equation $\mathbf{Ric} = 0$ then implies $M_{\mathrm{ADM}} = M_{\mathrm{Komar}} = 0$ which implies that spacetime is flat Minkowski. This theorem was originally shown for static spacetimes (i.e. hypersurface orthogonal K) by Pauli and Einstein [56] and later generalized to the stationary case by Lichnerowicz [97]. The result of this theorem cannot be circumvented by trying more complicated topologies for Σ. As soon as Σ becomes non-simply connected (which in view of the validity of the Poincaré conjecture will be the case for any one-ended manifold other than $\mathbb{R}^3$) we know from Gannon's theorem [63] that the evolving spacetime will inevitably develop singularities.

Finally we mention that under suitable fall-off conditions we can find the Poincaré group as *asymptotic symmetry group* [29]. It will emerge from (260) as equivalence classes of all hypersurface deformations, including those in which α and β asymptotically approach rigid translations, rotations, or boosts. The quotient is taken with respect to those deformations which are generated by the constraints, in which α and β tend to zero at spatial infinity. There are various subtleties and fine tunings involved for the precise fall-off conditions that are necessary in order to exactly obtain a 10-dimensional symmetry as a quotient of two infinite-dimensional objects. This is particularly true for asymptotic boosts, for which one needs to tilt the hypersurface, corresponding to asymptotic lapse functions $\alpha \propto r$. (Boosted hypersurfaces are known to exist in the development of asymptotically flat initial data [42].) But leaving the analytic details aside, the qualitative picture is quite generic for gauge field theories with long-ranging field configurations [69]: A proper physical symmetry group arises as quotient of a general covariance group with respect to a proper normal subgroup, the latter being defined to be that object that is generated by the constraints.

10   Black-Hole data

In this section we discuss some simple solutions to the vacuum Einstein equations without cosmological constant. We first specify to the simplest case of time symmetric conformally flat data. Time symmetry means that the initial extrinsic curvature vanishes, $K = 0$. The corresponding Cauchy surface will then be totally geodesic in the spacetime that emerges from it. The vector constraint (144b) is identically satisfied and the scalar constraint (144a) reduces to scalar flatness

$$R(h) = \mathbf{Scal}^D = 0 \,. \qquad (309)$$

Conformal flatness means that

$$h = \Omega^4 \, \delta \,, \qquad (310)$$

where δ is the flat metric. From (125a) we infer that (309) is equivalent to Ω being harmonic

$$\Delta_\delta \Omega = 0 \qquad (311)$$

where Δ_δ is the Laplacian with respect to the flat metric δ. We seek solutions Ω which are asymptotically flat for $r \to \infty$ and give rise to complete manifolds in the metric structure defined by g. The only spherically symmetric such solution is

$$\Omega(r) = 1 + \frac{r_0}{r} \,, \qquad (312)$$

where the integration constant r_0 can be related to the ADM mass (290) by

$$M_{\text{ADM}} = 2c^2 r_0/G \,. \tag{313}$$

This solution is defined on $\Sigma = \mathbb{R}^3 - \{0\}$. The metric on Σ so obtained is

$$h = \left(1 + \frac{r_0}{r}\right)^4 \left(dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)\right) . \tag{314}$$

It admits the following isometries

$$I_1(r,\theta,\varphi) := (r_0^2/r, \theta, \varphi) \,, \tag{315a}$$

$$I_2(r,\theta,\varphi) := (r_0^2/r, \pi - \theta, \varphi + \pi) \,. \tag{315b}$$

Note that the second is just a composition of the first with the antipodal map $(r,\theta,\varphi) \mapsto (r, \pi - \theta, \varphi + \pi)$ which is well defined on $\mathbb{R}^3 - \{0\}$. This makes I_2 a fixed-point free action. The fixed-point set of I_1 is the 2-sphere $r = r_0$. Note that generally a submanifold that is the fixed-point set of an isometry is necessarily totally geodesic (has vanishing extrinsic curvature). To see this, consider a geodesic that starts on and tangentially to this submanifold. Such a geodesic cannot leave the submanifold, for if it did we could use the isometry to map it to a different geodesic with identical initial conditions, in contradiction to the uniqueness of solutions for the geodesic equation. Hence the 2-sphere $r = r_0$ has vanishing extrinsic curvature and is therefore, in particular, a minimal surface (has vanishing trace of the extrinsic curvature). The geometry inside the sphere $r = r_0$ is isometric to that outside it. This is depicted in Fig. 4.

For the data $(h = (314), K = 0)$ on $\Sigma = \mathbb{R}^3 - \{0\}$ we actually know its maximal time evolution: It is the Kruskal spacetime [95][79] which maximally extends the exterior Schwarzschild spacetime. Figure 5 shows a conformal diagram of Kruskal spacetime.

In Kruskal coordinates (Kruskal [95] uses (v, u), Hawking Ellis [79] (t', x') for what we call (T, X)) (T, X, θ, φ), where T and X each range in $(-\infty, \infty)$ obeying $T^2 - X^2 < 1$, the Kruskal metric reads (as usual, we write $d\Omega^2$ for $d\theta^2 + \sin^2\theta\, d\varphi^2$):

$$g = \frac{8r_0^2}{r} \exp(-r/r_0) \left(-dT^2 + dX^2\right) + r^2 d\Omega^2 \,, \tag{316}$$
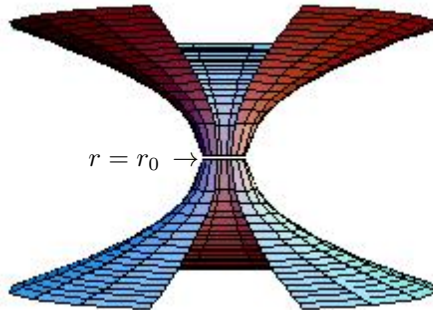
Figure 4: Cauchy surface with time symmetric initial data and two isometric asymptotically flat ends separated by a totally geodesic 2-sphere.

where r is a function of T and X, implicitly defined by

$$((r/r_0) - 1) \exp(r/r_0) = X^2 - T^2 \,. \tag{317}$$

The metric is spherically symmetric and allows for the additional Killing field

$$K = \left(X\partial_T + T\partial_X\right), \tag{318}$$

which is timelike for $|X| > |T|$ and spacelike for $|X| < |T|$.

Both maps (315) extend to the Kruskal manifold. The fixed-point free action (315b) has the extension

$$J : (T, X, \theta, \varphi) \mapsto (T, -X, \pi - \theta, \varphi + \pi) \,. \tag{319}$$

It generates a freely acting group $\mathbb{Z}_2$ of smooth isometries which preserve space- as well as time-orientation. Hence the quotient is a smooth space- and time-orientable manifold, that is sometimes called the $\mathbb{R}\mathrm{P}^3$-*geon*. It represents the maximal time evolution of the data $(h = (314), K = 0)$ as above, but now defined on the initial quotient manifold $\Sigma = (\mathbb{R}^3 - \{0\})/I_2$. It has only one asymptotically flat end and the topology of a once punctured real projective space $\mathbb{R}\mathrm{P}^3$. Note that the map J preserves the Killing field (318) only up to sign. Had one chosen $J' : (T, X, \theta, \varphi) \mapsto (-T, -X, \pi - \theta, \varphi + \pi)$ as in [101] and [66], one would have preserved K but lost time orientability.
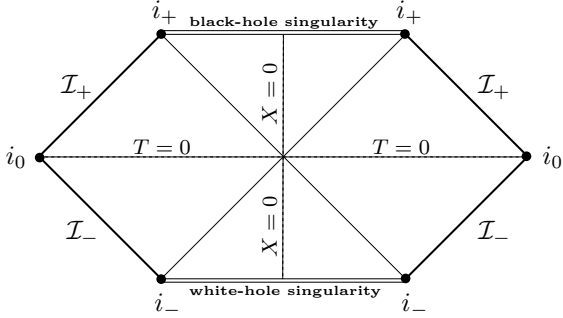
Figure 5: Conformally compactified Kruskal spacetime. The T axis points up vertically, the X axis horizontally to the right. The Cauchy surface of Fig. 4 corresponds trio the hypersurface $T = 0$. The various infinities are: i_0 spacelike, $i_\pm$ future/past timelike, and $\mathcal{I}_\pm$ future/past lightlike infinity. The right diamond-shaped region corresponds to the usual exterior Schwarzschild solution containing one asymptotically flat end.

Within the set of conformally-flat and time symmetric initial data we can easily generalize the solution (312) to (311) to include more than one monopole term on a multi-punctured $\mathbb{R}^3$. For two terms we get

$$\Omega(r) = 1 + \frac{a_1}{r_1} + \frac{a_2}{r_2}\,, \qquad (320)$$

where $r_i = \|\vec{x} - \vec{c}_i\|$. This represents two black holes without spin and orbital angular momentum momentarily at rest, with $\vec{c}_i \in \mathbb{R}^3$ representing the hole's "positions". The manifold has three ends, one for $r \to \infty$ and one each for $r_i \to 0$. For each end we can calculate the ADM mass and get

$$M = 2(a_1 + a_2)c^2/G\,, \qquad (321a)$$

$$M_1 = 2\left(a_1 + \frac{a_1 a_2}{r_{12}}\right)\frac{c^2}{G}\,, \qquad (321b)$$

$$M_2 = 2\left(a_2 + \frac{a_1 a_2}{r_{12}}\right)\frac{c^2}{G}\,, \qquad (321c)$$

where $r_{12} := \|\vec{c}_1 - \vec{c}_2\|$. Here M is the total mass associated with the end $r \to \infty$ and M_i is the individual hole mass associated with the end $r_i \to 0$. The binding energy is the overall energy minus the individual ones. One obtains

$$\Delta E := (M - M_1 - M_2)c^2 = -G\frac{M_1 M_2}{r_{12}} + \cdots \quad (322)$$

where the dots stand for corrections of quadratic and higher powers in $GM_i/c^2 r_{12}$. This can be easily generalized to any finite number of poles. Note that the initial manifolds are all complete, i.e. all punctures lie at infinite metric distance from any interior point.

Other generalizations consist in adding linear and angular momentum. This can be done using the *conformal method*, which we now briefly describe. Recall that we wish to solve the constraints (144) for $\mathbf{T} = 0$ but now with $K \neq 0$. Encouraged by previous experience with the simplifying effect of conformal transformations, we now study the general conformal transformation properties of the left-hand sides of (144). Generalizing (310), we write

$$h_{ab} = \Omega^4 \bar{h}_{ab}\,, \qquad (323a)$$

$$K^{ab} = \Omega^{-s}\bar{K}^{ab}\,. \qquad (323b)$$

Note that in view of (323a) the second equation is equivalent to

$$K_{ab} = \Omega^{8-s}\bar{K}_{ab}\,. \qquad (323c)$$

We first wish to determine the power s that is most suitable for simplifying (144b) if written in terms of $\bar{h}$ and $\bar{K}$. A slightly lengthy but straightforward computation gives

$$\begin{aligned} D_a\big(K^{ab} - h^{ab}K_c^c\big) = &\ \Omega^{-s}\bar{D}_a\big(\bar{K}^{ab} - \bar{h}^{ab}\bar{K}_c^c\big) \\ &+ (10 - s)\Omega^{-(s+1)}(\bar{D}_a\Omega)\bar{K}^{ab} \\ &+ (s - 6)\Omega^{-(s+1)}\bar{D}^b\bar{K}_c^c\,. \end{aligned} \qquad (324)$$

Here $\bar{D}$ is the Levi-Civita covariant derivative with respect to $\bar{h}$ and indices on barred quantities are moved with the barred metric. A suitable simplification in the sense of conformal covariance would occur if only the first line in (324) survived. The other two lines cannot be made to vanish simultaneously on account of a suitable choice of s. The best one can do is to choose $s = 10$ and restrict to traceless $\bar{K}$, i.e. $\bar{K}_c^c = 0$. From (323) this might seem as if we had to restrict to traceless K. But note that as the left-hand side of (144b) is linear

56

in K we can always add to any traceless solution $K^{(1)}$ a pure trace part

$$K^{(2)}_{ab} = \tfrac{1}{3}\tau h_{ab}\,, \qquad (325)$$

which satisfies the vector constraint as long as τ is constant. Putting all this together we see that we get a solution to the vector constraint if we maintain (323a) but replace (323b) with

$$K^{ab} = \Omega^{-10}\bar{K}^{ab} + \tfrac{1}{3}\Omega^{-4}\bar{h}^{ab}\tau\,, \qquad (326a)$$

$$K_{ab} = \Omega^{-2}\bar{K}_{ab} + \tfrac{1}{3}\Omega^{4}\bar{h}_{ab}\tau\,, \qquad (326b)$$

where τ is a constant and $\bar{K}$ is transverse traceless in the metric $\bar{h}$:

$$\bar{h}_{ab}\bar{K}^{ab} = 0\,, \qquad (327a)$$

$$\bar{D}_a\bar{K}^{ab} = 0\,. \qquad (327b)$$

Note that $K^a_a = \tau$ so that the method as presented here only produces initial data of constant mean curvature. It can be generalized to non-constant τ, see e.g. [86].

As before, the idea is now to let the remaining scalar constraint determine the conformal factor. Inserting (323a) and (326) into the scalar constraint (144a) and using (125), we obtain the following elliptic *York equation* for Ω

$$-\varepsilon\Big(\Delta_{\bar{h}} - \tfrac{1}{8}\mathbf{Scal}^{\bar{D}}\Big)\Omega + \tfrac{1}{8}\Omega^{-7}\bar{K}_{ab}\bar{K}^{ab} - \tfrac{1}{12}\Omega^5\tau^2 = 0\,. \qquad (328)$$

Existence and uniqueness of this equation for the Lorentzian case $\varepsilon = -1$ is discussed in the survey [86]. It may be further simplified if, as before, we assume conformally flat intitial data, i.e.

$$\bar{h} = \delta = \text{flat metric}\,. \qquad (329)$$

Then

$$-\varepsilon\Delta_\delta\Omega + \tfrac{1}{8}\Omega^{-7}\bar{K}_{ab}\bar{K}^{ab} - \tfrac{1}{12}\Omega^5\tau^2 = 0\,. \qquad (330)$$

where $\bar{K}$ is now transverse-traceless with respect to the flat connection (partial derivatives in suitable coordinates).

It is remarkable that the ADM momenta (291) can be calculated without knowing Ω. Hence we can parametrize solutions to (327) directly by the momenta without solving (328) first. Two solutions of particular interest for $\bar{h} = \delta$ are the *Bowen-York data* [37][119]. In Cartesian coordinates and corresponding components they read

$$\bar{K}^{(1)}_{ab} = r^{-2}\big(\nu_a A_b + \nu_b A_a - (\delta_{ab} - \nu_a\nu_b)\nu^c A_c\big)\,, \qquad (331a)$$

$$\bar{K}^{(2)}_{ab} = r^{-3}\big(\nu_a\varepsilon_{bcd} + \nu_b\varepsilon_{acd}\big)B^c\nu^d\,, \qquad (331b)$$

where $\nu^a := x^a/r$ and where all indices are raised and lowered with the flat metric δ. A and B are covariantly constant vector fields with respect to the Levi Civita connection for the flat metric δ, which are here represented by constant components A^b and B^c. One verifies by direct computation that they satisfy (327) with $\bar{h}_{ab} = \delta_{ab}$ and $\bar{D}_a = \partial_a$. Furthermore, using (291) one shows that (331a) has vanishing angular momentum and a linear momentum with components

$$P^a = \frac{2c^3}{3G}A^a\,, \qquad (332)$$

whereas (331b) has vanishing linear momentum and an angular momentum with components

$$J^a = \frac{c^3}{3G}B^a\,. \qquad (333)$$

They can be combined to give data for single holes with non-zero linear and angular momenta and also be superposed in order to give data for multi black-hole configurations. Such data, and certain modifications of them, form the essential ingredient for present-day numerical simulations of black hole scattering and the subsequent emission of gravitational radiation.

Let us now return to equation (328), which we have to solve once suitable expressions for the components $\bar{K}_{ab}$ have been found. As the metric is flat at the end representing spatial infinity, i.e. $\Omega(r \to \infty) = 1$, it is clear that Ω cannot be bounded in the interior region. The idea of the *puncture method*, first proposed in [38], is to restrict the type of singularities of Ω to be, in some sense, as simple as possible, which here means to

be of the pure monopole type that we already encountered in (312) for a single hole and generalized to two holes in (320). This amounts to the following: Take $\Sigma = \mathbb{R}^3 - \{\vec{c}_1, \cdots, \vec{c}_n\}$ and assume we are given suitable $\bar{K}_{ab}$ which are regular in Σ, e.g., a sum of York data of the form (331), each centered at one of the punctures $\vec{c}_i$. Accordingly, we write

$$\Omega = u + \frac{1}{\omega}, \qquad (334a)$$

where

$$\frac{1}{\omega} := \sum_{i=1}^{n} \frac{a_i}{r_i}, \qquad (334b)$$

again with $r_i := \|\vec{x} - \vec{c}_i\|$. The crucial assumption now is that the function u is smooth, at least C^2, on *all* of $\mathbb{R}^3$, *including* the points $\vec{c}_i$. Since $1/\omega$ is annihilated by Δ_δ, (328) applied to (334) then leads to a second order elliptic differential equation for u. In the simple case of conformally flat maximal data, i.e. $\bar{h}_{ab} = \delta_{ab}$ and $\tau = 0$, we get in the Lorentzian case ($\varepsilon = -1$)

$$\Delta_\delta u = -\tfrac{1}{8}\left(1 + \omega u\right)^{-7} \omega^7 \bar{K}_{ab} \bar{K}^{ab}, \qquad (335)$$

where $u \to 1$ at spatial infinity. Now, at the i-th puncture, $\bar{K}_{ab}$ diverges as $(1/r_i)^2$ for the data (331a) and as $(1/r_i)^3$ for the data (331b). This means that $\bar{K}_{ab}\bar{K}^{ab}$ diverges at most as $(1/r_i)^6$. But from (334b) we see that ω vanishes as r_i at $\vec{c}_i$ so that $\omega^7 \bar{K}_{ab}\bar{K}^{ab}$ also vanishes at least as fast as r_i at $\vec{c}_i$ and is hence continuous on all of $\mathbb{R}^3$. Standard elliptic theory now allows to conclude existence and uniqueness of C^2 solutions to (335); see [38] for more details. It is then not difficult to see that the Riemannian manifold (Σ, h) so obtained has $n + 1$ asymptotically flat ends whose ADM masses are readily calculated. Similar to (321), one obtains

$$M = \frac{2c^2}{G} \sum_{i=1}^{n} a_i, \qquad (336a)$$

$$M_i = \frac{2c^2}{G}\left(a_i u_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_i a_j}{r_{ij}}\right), \qquad (336b)$$

where $u_i := u(\vec{c}_i)$ and $r_{ij} := \|\vec{c}_i - \vec{c}_j\|$. Comparing this to (321) (and its obvious generalization from

2 to n punctures) shows that the only difference in the analytic expression for the masses is the appearance of u_i (instead of 1) in the first term on the right-hand side of (336b). Thus, formally, for fixed monopole parameters a_i, the switching-on of linear and angular momentum (here represented locally by trace-free extrinsic curvatures) adds to each individual mass a term $a_i(u_i - 1)$. This contribution is non-negative as a consequence of $u \geq 1$. The latter equation follows immediately from standard elliptic theory. Indeed, (335) implies $\Delta_\delta u \leq 0$, i.e. that u is *superharmonic*, and hence that $u \geq f$ for any continuous harmonic f with the same boundary values, i.e. $f \equiv 1$. Note that this argument relies on $\varepsilon = -1$.

Finally we wish to point out an interesting type on non-uniqueness in writing down initial data of the form (331). It has to do with the question of whether we wish to enforce the inversion symmetries of the type (315) to become isometries of the initial geometry. Let us focus on I_1 as defined in (315), where we now denote the radius of inversion a (rather than r_0). Dropping the subscript 1, we have

$$[I(x)]^a = (a^2/r^2)\, x^a. \qquad (337)$$

Its Jacobian is

$$[I_*(x)]^a_b = (a^2/r^2)\left(\delta^a_b - 2\nu^a \nu_b\right). \qquad (338)$$

We note in passing that the matrix in round brackets is orthogonal with determinant -1.

It follows that the conformally flat metric $h = \Omega^4 \delta$ satisfies $I^* h = h$ iff

$$(a/r)(\Omega \circ I) = \Omega. \qquad (339)$$

In such a metric the sphere $r = a$ is the fixed-point set of the isometry I and hence totally geodesic. In particular, this implies that it is a stationary point of the area function which is equivalent to $\partial_r(r^2 \Omega^4) = 0$ and hence to

$$\left[\frac{\partial \Omega}{\partial r} + \frac{\Omega}{2a}\right]_{r=a} = 0, \qquad (340)$$

which may also be directly verified from differentiating (339) with respect to r at $r = a$. This would

be the condition for (330) at the "inner boundary" in order to produce a solution that gives rise to a metric h that has I as an isometry. But that, clearly, also puts conditions on the extrinsic curvature $\bar{K}$, for h and K have to satisfy the coupled system of constraints (144) in the vacuum case $\mathbf{T} = 0$. A sufficient condition is

$$I^* K = \pm K \,. \tag{341a}$$

Using (326b) and restricting to maximal data, $K_c^c = \tau = 0$, this is equivalent to

$$(a^2/r^2)\, I^* \bar{K} = \pm \bar{K} \,, \tag{341b}$$

where we also made use of (339).

Using the global chart $\{x^1, x^2, x^3\}$ on $\mathbb{R}^3 - \{0\}$ together wit the flat metric $\bar{h} = \delta$, we have the function r whose value at x is the δ-geodesic distance $((x^1)^2 + (x^2)^2 + (x^3)^2)$ and the following vector fields and volume form (A^a, B^a, ϵ_{abc} being constant component functions)

$$\nu = \frac{x^a}{r} \frac{\partial}{\partial x^a} \,, \tag{342a}$$

$$A = A^a \frac{\partial}{\partial x^a} \,, \tag{342b}$$

$$B = B^a \frac{\partial}{\partial x^a} \,, \tag{342c}$$

$$\varepsilon = \tfrac{1}{3!} \varepsilon_{abc} dx^a \wedge dx^b \wedge dx^c \,. \tag{342d}$$

The co-vector fields that arise from these vector fields via the isomorphism induced by the flat metric δ are called $\nu^\flat, A^\flat, B^\flat$. Under the inversion-map (337), making also use of (338), these structures behave as follows:

$$I^* r = r \circ I = \frac{a^2}{r} \,, \tag{343}$$

$$I_* \nu = -(a/r)^{-2} \nu \,, \tag{344a}$$

$$I^* \nu^\flat = -(a/r)^2 \nu^\flat \,, \tag{344b}$$

$$I_* A = (a/r)^{-2} \big(A - 2\,\nu(\nu \cdot A) \big) \,, \tag{345a}$$

$$I^* A^\flat = (a/r)^2 \big(A^\flat - 2\,\nu^\flat(\nu \cdot A) \big) \,, \tag{345b}$$

and identically for B, where a dot denotes the scalar product with respect to the flat metric, i.e. $\nu \cdot A = \delta(\nu, A)$,

$$I^* \delta = (a/r)^4\, \delta \,, \tag{346}$$

and[2]

$$I^* \varepsilon = -(a/r)^6 \varepsilon \,. \tag{347}$$

These formulae allow to immediatly write down the I-transforms of the data (331). We have, in components,

$$I^* \bar{K}^{(1)}_{ab}$$
$$= -r^{-2} \big(\nu_a A_b + \nu_b A_a + (\delta_{ab} - 5\,\nu_a\nu_b)\nu^c A_c \big). \tag{348a}$$

and

$$I^* \bar{K}^{(2)}_{ab} = -(a/r)^{-2} \bar{K}^{(2)} \,. \tag{348b}$$

This means that $\bar{K}^{(2)}$ as given by (331b) is already antisymmetric in the sense of (341b), but (331a) is neither symmetric nor antisymmetric. Symmetric or antisymmetric data can be obtained by forming the symmetric or antisymmetric combination

$$\bar{K}^{(1)}_\pm := \bar{K}^{(1)} \pm (a/r)^2 I^* \bar{K}^{(1)} \,. \tag{349}$$

which satisfies

$$(a/r)^2 I^* \bar{K}^{(1)}_\pm = \pm \bar{K}^{(1)}_\pm \,. \tag{350}$$

In components they read

$$\big[\bar{K}^{(1)}_\pm\big]_{ab}$$
$$= 1/r^2 \big(\nu_a A_b + \nu_b A_a - (\delta_{ab} - \nu_a\nu_b)\nu^c A_c \big)$$
$$\mp a^2/r^4 \big(\nu_a A_b + \nu_b A_a + (\delta_{ab} - 5\,\nu_a\nu_b)\nu^c A_c \big). \tag{351}$$

Having enforced symmetry with respect to the inversion (337) we will obtain an initial-data 3-manifold with two isometric asymptotically flat

---

[2] A straightforward calculation using (338) first yields $I^* \varepsilon = (a/r)^6 \big(\varepsilon - 2\,\nu^\flat \wedge \star\nu^\flat\big)$, where $\star$ denotes the Hodge dual with respect to δ. But for any vector field ν one trivially has $\nu^\flat \wedge \varepsilon = 0$ and hence, now assuming ν to be also normalized, $0 = i_\nu(\nu^\flat \wedge \varepsilon) = \varepsilon - \nu^\flat \wedge i_\nu\varepsilon$. Using $i_\nu\varepsilon = \star\nu^\flat$, this gives $\nu^\flat \wedge \star\nu^\flat = \varepsilon$ and hence (347).

ends whose Poincaré charges coincide (possibly up to sign) and are given by those of the original data. This is trivially true for angular momentum and follows for linear momentum from the $1/r^4$ fall-off of the second term in (351). It is interesting to note that a term proportional to the second term in (351) follows in case of spherically symmetric extended matter sources [36].

11 Further developments, problems, and outlook

In this contribution we have explained in some detail the dynamical and Hamiltonian formulation of GR. We followed the traditional ADM approach in which the basic variables are the Riemannian metric h of space and its conjugate momentum π, which is essentially the extrinsic curvature that Σ will assume once the spacetime is developed and Σ is isometrically embedded in it. Attempts to establish a theory of *Quantum Gravity* based on the Hamiltonian formulation of GR suggest that other canonical variables are better suited for the mathematical implementation of the constraints and the ensuing construction of spaces of states and observables [113][107][34]. These variables are a (suitably densitized) orthonormal 3-bein field E on Σ and the *Ashtekar-Barbero connection*. We have already seen that orientable Σ are parallelizable so that global fields E do indeed exist. Any field E determines a Riemannian metric h, which in turn determines its Levi-Civita connection. The Ashtekar-Barbero covariant derivative, $\mathcal{D}$, differs from the Levi-Civita connection D of h by the endomorphism-valued 1-form which associates to each tangent vector X the tangent-space endomorphism $Y \mapsto \gamma \mathbf{Wein}(X) \times Y$, where γ is a dimensionless constant, the so-called Barbero-Immirzi-parameter, which was first introduced by Immirzi in [85] on the basis of Barbero's generalization [22] of Ashtekar's variables. Hence we have

$$\mathcal{D}_X Y = D_X Y + \gamma \mathbf{Wein}(X) \times Y. \qquad (352)$$

The multiplication $\times$ is the standard 3-dimensional vector product with respect to the metric h. It is

defined as follows

$$X \times Y := \left[\star(X^\flat \wedge Y^\flat) \right]^\sharp, \qquad (353)$$

where the isomorphisms $\flat$ and $\sharp$ are with respect to h (cf. (1)). The product $\times$ obeys the standard rules: It is bilinear, antisymmetric, and $X \times (Y \times Z) = h(X, Z)Y - h(X, Y)Z$. Moreover, for any X, the endomorphism $Y \mapsto X \times Y$ is antisymmetric with respect to h, i.e. $h(X \times Y, Z) = -h(Y, X \times Z)$, and hence it is in the Lie algebra of the orthogonal group of h. In particular this is true for $Y \mapsto \mathbf{Wein}(X) \times Y$, showing that $\mathcal{D}$ is again metric, i.e. obeys $\mathcal{D}h = 0$ once its unique extension to all tensor fields is understood. Clearly, unlike D, the torsion of $\mathcal{D}$ cannot be zero:

$$\begin{aligned} T^{\mathcal{D}}(X, Y) &= \mathcal{D}_X Y - \mathcal{D}_Y X - [X, Y] \\ &= \gamma \big(\mathbf{Wein}(X) \times Y - \mathbf{Wein}(Y) \times X \big). \end{aligned} \qquad (354)$$

Using (77) and index notation, the curvature tensor for $\mathcal{D}$ is

$$\begin{aligned} R^{\mathcal{D}}_{abcd} &= R^D_{abcd} \\ &\quad + \varepsilon\gamma \big(D_c K_{dn} - D_d K_{cn} \big) \varepsilon^n{}_{ab} \\ &\quad - \gamma^2 \big(K_{ac} K_{bd} - K_{ad} K_{bc} \big). \end{aligned} \qquad (355)$$

From this the scalar curvature follows

$$\mathbf{Scal}^{\mathcal{D}} = \mathbf{Scal}^D + \gamma^2 \, G^{abcd} K_{ab} K_{cd}. \qquad (356)$$

Camparison with (144a) shows that for $\gamma^2 = \varepsilon$ the gravitational part of the scalar constraint is just (ε times) the scalar curvature of $\mathcal{D}$. This striking simplification of the scalar constraint formed the original motivation for the introduction of $\mathcal{D}$ by Ashtekar [18]. However, for $\varepsilon = -1$ one needs to complexify the tensor bundle over Σ for $\gamma = \pm i$ to make sense, and subsequently impose reality conditions which re-introduce a certain degree of complication; see, e.g., [68] for a compact account not using spinors. The usage of $\mathcal{D}$ in the real case was then proposed by Barbero in [22] and forms the basic tool in *Loop Quantum Gravity* [113], which has definite technical advantages over the metric-based traditional approach.

At this point we wish to inject one word of caution concerning the possible geometric interpretation of the of Barbero connection $\mathcal{D}$, depending on the value of γ. From the defining equation (352) it is clear that any $\mathcal{D}$ explicitly contains extrinsic information, i.e. information that refers to the way Σ is embedded into spacetime M. Moreover, unless $\gamma^2 = \varepsilon$, this dependence on the embedding is such that $\mathcal{D}$ cannot be considered as pull-back of a connection defined on (the bundle of linear frames over) M, as has been pointed out in [109]. The argument is simple: If it were the pull-back of such a connection on M, its holonomy along a loop in spacetime would be the same for all Σ containing that loop. That this is not the case for $\gamma^2 \neq \varepsilon$ can, e.g., be checked for the simple example where M is Minkowski space and the loop is a planar unit circle that is contained in a flat spacelike hyperplane as well as in the constant-curvature spacelike hyperboloid of unit future-pointing timelike vectors. The holonomy in the latter case turns out to be non-trivial (see [109] for the explicit calculation). For the Ashtekar connection, i.e. for $\gamma^2 = \varepsilon$, we know from its original construction that it is the pull-back of a spacetime connection (see, e.g., the derivation in [68]). But for Barbero's generalizations (352) with $\gamma^2 \neq \varepsilon$, and in particular for all real values of γ in the Lorentzian case ($\varepsilon = -1$), this means that it is impossible to attach a gauge-theoretic *spacetime* interpretation to $\mathcal{D}$.

Despite this conceptual shortcoming, the technical advantages over the metric-based approach remain. On the other hand, the latter is well suited to address certain conceptual problems [92], like e.g. the problem of time that emerges in those cases where the Hamiltonian (241) has no boundary terms and is therefore just a sum of constraints. This happens in cosmology based on closed Σ. The motions generated by the Hamiltonian are then just pure gauge transformations and the question arises whether and how 'motion' and 'change' are to be recovered; see, e.g., [107, 108].

Dynamical models in cosmology often start from symmetry assumptions that initially reduce the infinitely many degrees of freedom to finitely many ones (so-called mini-superspace models). Other modes are then treated perturbatively in an expansion around the symmetric configurations. In these cases quantization in the metric representation can be performed, with potentially interesting consequences for observational cosmology, like the modification of the anisotropy spectrum of the cosmic microwave background [93][32]. All these attempts make essential use of the Hamiltonian theory as described in this contribution.

12 Appendix: Group actions on manifolds

Let G be a group and M a set. An *action* of G on M is a map

$$\Phi : G \times M \to M \qquad (357)$$

such that, for all $m \in M$ and $e \in G$ the neutral element,

$$\Phi(e, m) = m \,, \qquad (358)$$

and where, in addition, one of the following two conditions holds:

$$\Phi\big(g, \Phi(h, m)\big) = \Phi(gh, m) \,, \qquad (359a)$$
$$\Phi\big(g, \Phi(h, m)\big) = \Phi(hg, m) \,. \qquad (359b)$$

If (357), (358), and (359a) hold we speak of a *left action*. A *right action* satisfies (357), (358), and (359b). For a left action we also write

$$\Phi(g, m) =: g \cdot m \qquad (360a)$$

and for a right action

$$\Phi(g, m) =: m \cdot g \,. \qquad (360b)$$

Equations (359) then simply become (group multiplication is denoted by juxtaposition without a dot)

$$g \cdot (h \cdot m) = (gh) \cdot m \,, \qquad (361a)$$
$$(m \cdot h) \cdot g = m \cdot (hg) \,. \qquad (361b)$$

Holding either of the two arguments of Φ fixed we obtain the families of maps

$$\begin{aligned} \Phi_g : \; & M \to M \\ & m \mapsto \Phi(g, m) \end{aligned} \qquad (362)$$

for each $g \in G$, or

$$\Phi_m : \begin{array}{l} G \to M \\ g \mapsto \Phi(g, m) \end{array} \qquad (363)$$

for each $m \in M$. Note that (358) and (359) imply that $\Phi_{g^{-1}} = (\Phi_g)^{-1}$. Hence each Φ_g is a bijection of M. The set of bijections of M will be denoted by $\mathrm{Bij}(M)$. It is naturally a group with group multiplication being given by composition of maps and the neutral element being given by the identity map. Conditions (358) and (359a) are then equivalent to the statement that the map $G \to \mathrm{Bij}(M)$, given by $g \mapsto \Phi_g$, is a group homomorphism. Likewise, (358) and (359b) is equivalent to the statement that this map is a group anti-homomorphism.

The following terminology is standard: The set $\mathrm{Stab}(m) := \{g \in G : \Phi(g, m) = m\} \subset G$ is called the *stabilizer* of m. It is easily proven to be a normal subgroup of G satisfying $\mathrm{Stab}(g \cdot m) = g\big(\mathrm{Stab}(g \cdot m)\big)g^{-1}$ for left and $\mathrm{Stab}(m \cdot g) = g^{-1}\big(\mathrm{Stab}(g \cdot m)\big)g$ for right actions. The *orbit* of G through $m \in M$ is the set $\mathrm{Orb}(m) := \{\Phi(g, m) : g \in G\} =: \Phi(G, m)$ (also written $G \cdot m$ for left and $m \cdot G$ for right action). It is easy to see that two orbits are either disjoint or identical. Hence the orbits partition M. A point $m \in M$ is called a *fixed point* of the action Φ iff $\mathrm{Stab}(m) = G$. An action Φ is called *effective* iff $\Phi(g, m) = m$ for all $m \in M$ implies $g = e$; i.e., "only the group identity moves nothing". Alternatively, we may say that effectiveness is equivalent to the map $G \mapsto \mathrm{Bij}(M)$, $g \mapsto \Phi_g$, being injective; i.e., $\Phi_g = \mathrm{Id}_M$ implies $g = e$. The action Φ is called *free* iff $\Phi(g, m) = m$ for some $m \in M$ implies $g = e$; i.e., "no $g \neq e$ fixes a point". This is equivalent to the injectivity of all maps $\Phi_m : G \to M$, $g \mapsto \Phi(g, m)$, which can be expressed by saying that all orbits of G in M are faithful images of G.

Here we are interested in *smooth* actions. For this we need to assume that G is a Lie group, that M is a differentiable manifold, and that the map (357) is smooth. We denote by $\exp : T_eG \to G$ the exponential map. For each $X \in T_eG$ there is a

vector field V^X on M, given by

$$\begin{aligned} V^X(m) &= \frac{d}{dt}\Big|_{t=0} \Phi\big(\exp(tX), m\big) \\ &= \Phi_{m*e}(X). \end{aligned} \qquad (364)$$

Recall that Φ_{m*e} denotes the differential of the map Φ_m evaluated at $e \in G$. V^X is also called the *fundamental vector field* on M associated to the action Φ of G and to $X \in T_eG$. (We will later write $\mathrm{Lie}(G)$ for T_eG, after we have discussed *which* Lie structure on T_eG we choose.)

In passing we note that from (364) it already follows that the flow map of V^X is given by

$$\mathrm{Fl}_t^{V^X}(m) = \Phi(\exp(tX), m). \qquad (365)$$

This follows from $\exp(sX)\exp(tX) = \exp\big((s+t)X\big)$ and (359) (any of them), which imply

$$\mathrm{Fl}_s^{V^X} \circ \mathrm{Fl}_t^{V^X} = \mathrm{Fl}_{s+t}^{V^X} \qquad (366)$$

on the domain of M where all three maps appearing in (366) are defined. Uniqueness of flow maps for vector fields then suffices to show that (365) is indeed the flow of V^X.

Before we continue with the general case, we have a closer look at the special cases where $M = G$ and Φ is either the left translation of G on G, $\Phi(g, h) = L_g(h) := gh$, or the right translation, $\Phi(g, h) = R_g(h) := hg$. The corresponding fundamental vector fields (364) are denoted by V_R^X and V_L^X respectively:

$$V_R^X(h) = \frac{d}{dt}\Big|_{t=0}\Big(\exp(tX)\,h\Big), \qquad (367a)$$

$$V_L^X(h) = \frac{d}{dt}\Big|_{t=0}\Big(h\,\exp(tX)\Big). \qquad (367b)$$

The seemingly paradoxical labeling of R for left and L for right translation finds its explanation in the fact that V_R^X is right and V_L^X is left invariant, i.e., $R_{g*}V_R^X = V_R^X$ and $L_{g*}V_L^X = V_L^X$. Recall that the latter two equations are shorthands for

$$R_{g*h}V_R^X(h) = V_R^X(hg), \qquad (368a)$$

$$L_{g*h}V_L^X(h) = V_L^X(gh). \qquad (368b)$$

The proofs of (368a) only uses (367a) and the chain rule:

$$
\begin{aligned}
R_{g*h}V_R^X(h) &= R_{g*h}\frac{d}{dt}\Big|_{t=0}\Big(\exp(tX)\,h\Big) \\
&= \frac{d}{dt}\Big|_{t=0}R_g\Big(\exp(tX)\,h\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\Big(\exp(tX)\,hg\Big) \\
&= V_R^X(hg)\,.
\end{aligned}
\tag{369a}
$$

Similarly, the proof of (368b) starts from (367b):

$$
\begin{aligned}
L_{g*h}V_L^X(h) &= L_{g*h}\frac{d}{dt}\Big|_{t=0}\Big(h\,\exp(tX)\Big) \\
&= \frac{d}{dt}\Big|_{t=0}L_g\Big(h\,\exp(tX)\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\Big(gh\,\exp(tX)\Big) \\
&= V_L^X(gh)\,.
\end{aligned}
\tag{369b}
$$

In particular, we have

$$
V_R^X(g) = R_{g*e}V_L^X(e) = R_{g*e}X\,, \tag{370a}
$$
$$
V_L^X(g) = L_{g*e}V_R^X(e) = L_{g*e}X\,, \tag{370b}
$$

showing that the vector spaces of right/left invariant vector fields on G are isomorphic to T_eG. Moreover, the vector spaces of right/left invariant vector fields on G are Lie algebras, the Lie product being their ordinary commutator (as vector fields). This is true because the operation of commuting vector fields commutes with push-forward maps of diffeomorphisms: $\phi_*[V,W] = [\phi_*V, \phi_*W]$. This implies that the commutator of right/left invariant vector fields is again right/left invariant. Hence the isomorphisms can be used to turn T_eG into a Lie algebra, identifying it either with the Lie algebra of right- or left-invariant vector fields. The standard convention is to choose the latter. Hence, for any $X, Y \in \mathrm{Lie}(G)$, one defines

$$
[X,Y] := [V_L^X, V_L^Y](e)\,. \tag{371}
$$

T_eG endowed with *that* structure is called $\mathrm{Lie}(G)$. Clearly, this turns $V_L : \mathrm{Lie}(G) \to \Gamma TG$, $X \mapsto V_L^X$, into a Lie homomorphism:

$$
V_L^{[X,Y]} = [V_L^X, V_L^Y]\,. \tag{372}
$$

As a consequence, $V_R : \mathrm{Lie}(G) \to \Gamma TG$, $X \mapsto V_R^X$, now turns out to be an *anti* Lie homomorphism, i.e., to contain an extra minus sign:

$$
V_R^{[X,Y]} := -[V_R^X, V_R^Y]\,. \tag{373}
$$

This can be proven directly but will also follow from the more general considerations below.

On G consider the map

$$
\begin{aligned}
C : G \times G &\to G \\
(h,g) &\mapsto hgh^{-1}\,.
\end{aligned}
\tag{374}
$$

For fixed h this map, $C_h : G \to G$, $g \mapsto C_h(g) = hgh^{-1}$, is an automorphism (i.e., self-isomorphism) of G. Automorphisms of G form a group (multiplication being composition of maps) which we denote by $\mathrm{Aut}(G)$. It is immediate that the map $C \to \mathrm{Aut}(G)$, $h \mapsto C_h$, is a homomorphism of groups; i.e.,

$$
C_e = \mathrm{Id}_G\,, \tag{375a}
$$
$$
C_h \circ C_k = C_{hk}\,. \tag{375b}
$$

Taking the differential at $e \in G$ of C_h we obtain a linear self-map of T_eG, which we call Ad_h:

$$
\mathrm{Ad}_h := C_{h*e} : T_eG \to T_eG\,. \tag{376a}
$$

Differentiating both sides of both equations (375) at $e \in G$, using the chain rule together with $C_k(e) = e$ for the second, we infer that

$$
\mathrm{Ad}_e = \mathrm{Id}_{T_eG}\,, \tag{376b}
$$
$$
\mathrm{Ad}_h \circ \mathrm{Ad}_k = \mathrm{Ad}_{hk}\,. \tag{376c}
$$

This implies, firstly, that each linear map (376a) is invertible, i.e. an element of the *general linear group* $\mathrm{GL}(T_eG)$ of the vector space T_eG, and, secondly, that the map

$$
\begin{aligned}
\mathrm{Ad} : G &\to \mathrm{GL}(T_eG) \\
h &\mapsto \mathrm{Ad}_h
\end{aligned}
\tag{377}
$$

is a group homomorphism. In other words, Ad is a linear representation of G on T_eG, called the *adjoint representation*.

In (368) we saw that V_R^X and V_L^X are invariant under the action of right and left translations respectively (hence their names). But what happens if we act on V_R^X with left and on V_L^X with right translations? The answer is obtained from straightforward computation. In the first case we get:

$$
\begin{aligned}
L_{g*h}\big(V_R^X(h)\big) &= L_{g*h}\frac{d}{dt}\Big|_{t=0}\big(\exp(tX)\,h\big) \\
&= \frac{d}{dt}\Big|_{t=0}\big(g\,\exp(tX)\,h\big) \\
&= \frac{d}{dt}\Big|_{t=0}\big(C_g(\exp(tX))\,gh\big) \\
&= V_R^{\mathrm{Ad}_g(X)}(gh)\,,
\end{aligned}
$$
$$(378\mathrm{a})$$

where we used (376) in the last and the definition of V_R^X in the first and last step. Similarly, in the second case we have

$$
\begin{aligned}
R_{g*h}\big(V_L^X(h)\big) &= R_{g*h}\frac{d}{dt}\Big|_{t=0}\big(h\,\exp(tX)\,h\big) \\
&= \frac{d}{dt}\Big|_{t=0}\big(h\,\exp(tX)\,g\big) \\
&= \frac{d}{dt}\Big|_{t=0}\big(hg\,C_{g^{-1}}(\exp(tX))\big) \\
&= V_L^{\mathrm{Ad}_{g^{-1}}(X)}(gh)\,.
\end{aligned}
$$
$$(378\mathrm{b})$$

Taking the differential of Ad at $e \in G$ we obtain a linear map from T_eG into $\mathrm{End}(T_eG)$, the linear space of endomorphisms of T_eG (linear self-maps of T_eG).

$$
\begin{aligned}
\mathrm{ad} := \mathrm{Ad}_{*e} : T_eG &\to \mathrm{End}(T_eG) \\
X &\mapsto \mathrm{ad}_X\,.
\end{aligned}
$$
$$(379)$$

Now, we have

$$\mathrm{ad}_X(Y) = [X,Y] \tag{380}$$

where the right-hand side is defined in (371). The proof of (380) starts from the fact that the commutator of two vector fields can be expressed in terms of the Lie derivative of the second with respect to the first vector field in the commutator, and the

definition of the Lie derivative. We recall from (365) that the flow of the left invariant vector fields is given by right translation: $\mathrm{Fl}_t^{V_L^X}(g) = g\,\exp(tX)$. Then we have

$$
\begin{aligned}
[X,Y] &= [V_L^X, V_L^Y](e) \\
&= (L_{V_L^X}V_L^Y)(e) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Fl}_{(-t)*}^{V_L^X}\Big(V_L^Y(\mathrm{Fl}_t^{V_L^X}(e))\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Fl}_{(-t)*}^{V_L^X}\frac{d}{ds}\Big|_{s=0}\mathrm{Fl}_s^{V_L^Y}\Big(\mathrm{Fl}_t^{V_L^X}(e)\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0}\exp(tX)\exp(sY)\exp(-tX) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Ad}_{\exp(tX)}(Y) \\
&= \mathrm{ad}_X(Y)\,.
\end{aligned}
$$
$$(381\mathrm{a})$$

A completely analogous consideration, now using $\mathrm{Fl}_t^{V_R^X}(g) = \exp(tX)\,g$, allows to compute the commutator of the right-invariant vector fields evaluated at $e \in G$:

$$
\begin{aligned}
[V_R^X, V_R^Y](e) &= (L_{V_R^X}V_R^Y)(e) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Fl}_{(-t)*}^{V_R^X}\Big(V_R^Y(\mathrm{Fl}_t^{V_R^X}(e))\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Fl}_{(-t)*}^{V_R^X}\frac{d}{ds}\Big|_{s=0}\mathrm{Fl}_s^{V_R^Y}\Big(\mathrm{Fl}_t^{V_R^X}(e)\Big) \\
&= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0}\exp(-tX)\exp(sY)\exp(tX) \\
&= \frac{d}{dt}\Big|_{t=0}\mathrm{Ad}_{\exp(-tX)}(Y) \\
&= -\mathrm{ad}_X(Y) \\
&= -[X,Y]\,.
\end{aligned}
$$
$$(381\mathrm{b})$$

Equation (373) now follows if we act on both sides of $[V_R^X, V_R^Y](e) = -[X,Y]$ with R_{g*e} and use (368a).

We now return to the general case where M is any manifold and the vector field V^X is defined by an action Φ as in (364) and whose flow map is given by (365). Now, given that Φ is a *right action*, we

64

obtain

$$\left[V^X, V^Y\right](m)$$
$$= (L_{V^X} V^Y)(m)$$
$$= \frac{d}{dt}\Big|_{t=0} \mathrm{Fl}^{V^X}_{(-t)*}\left(V^Y(\mathrm{Fl}^{V^X}_t(m))\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \mathrm{Fl}^{V^X}_{(-t)*}\frac{d}{ds}\Big|_{s=0} \mathrm{Fl}^{V^Y}_s\left(\mathrm{Fl}^{V^X}_t(m)\right)$$
$$= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0} \Phi\left(\exp(tX)\exp(sY)\exp(-tX), m\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi_{m*e}\left(\mathrm{Ad}_{\exp(tX)}(Y)\right)$$
$$= V^{\mathrm{ad}_X(Y)}(m)$$
$$= V^{[X,Y]}(m)$$

$$(382a)$$

where we used (365) and (359b) at the fourth and (380) at the last equality. Similarly, if Φ is a *left action*, we have

$$\left[V^X, V^Y\right](m)$$
$$= (L_{V^X} V^Y)(m)$$
$$= \frac{d}{dt}\Big|_{t=0} \mathrm{Fl}^{V^X}_{(-t)*}\left(V^Y(\mathrm{Fl}^{V^X}_t(m))\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \mathrm{Fl}^{V^X}_{(-t)*}\frac{d}{ds}\Big|_{s=0} \mathrm{Fl}^{V^Y}_s\left(\mathrm{Fl}^{V^X}_t(m)\right)$$
$$= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0} \Phi\left(\exp(-tX)\exp(sY)\exp(tX), m\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi_{m*e}\left(\mathrm{Ad}_{\exp(-tX)}(Y)\right)$$
$$= -V^{\mathrm{ad}_X(Y)}(m)$$
$$= -V^{[X,Y]}(m)$$

$$(382b)$$

where we used (365) and (359a) at the fourth and again (380) at the last equality.

Finally we derive the analog of (378) in the general case. This corresponds to computing the pushforward of V^X under Φ_g. If Φ is a left action we will obtain the analog of (378a), and the analog of (378b) if Φ is a right action. For easier readability we shall also make use of the notation (360). For

a left action we then get

$$\Phi_{g*m}\left(V^X(m)\right) = \Phi_{g*m}\frac{d}{dt}\Big|_{t=0} \Phi(\exp(tX), m)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi\left(g\,\exp(tX), m\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi\left(C_g(\exp(tX)), g\cdot m\right)$$
$$= \Phi_{(g\cdot m)*e}\frac{d}{dt}\Big|_{t=0} C_g\left(\exp(tX)\right)$$
$$= \Phi_{(g\cdot m)*e}\left(\mathrm{Ad}_g(X)\right)$$
$$= V^{\mathrm{Ad}_g(X)}(g\cdot m)$$
$$= V^{\mathrm{Ad}_g(X)}\left(\Phi(g,m)\right).$$

$$(383a)$$

Similarly, if Φ is a right action,

$$\Phi_{g*m}\left(V^X(m)\right) = \Phi_{g*m}\frac{d}{dt}\Big|_{t=0} \Phi(\exp(tX), m)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi\left(\exp(tX)\,g, m\right)$$
$$= \frac{d}{dt}\Big|_{t=0} \Phi\left(C_{g^{-1}}(\exp(tX)), m\cdot g\right)$$
$$= \Phi_{(m\cdot g)*e}\frac{d}{dt}\Big|_{t=0} C_{g^{-1}}\left(\exp(tX)\right)$$
$$= \Phi_{(m\cdot g)*e}\left(\mathrm{Ad}_{g^{-1}}(X)\right)$$
$$= V^{\mathrm{Ad}_{g^{-1}}(X)}(m\cdot g)$$
$$= V^{\mathrm{Ad}_{g^{-1}}(X)}\left(\Phi(g,m)\right).$$

$$(383b)$$

References

[1] Marcus Ansorg and David Petroff. Negative Komar mass of single objects in regular, asymptotically flat spacetimes. *Classical and Quantum Gravity*, 23(24):L81–L87, 2006.

[2] Vladimir Igorevich Arnold. *Mathematical Methods of Classical Mechanics*. Springer Verlag, New York, 1978.

[3] Richard Arnowitt and Stanley Deser. Quantum theory of gravitation: General formulation and linearized theory. *Physical Review*, 113(2):745–750, 1959.

[4] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Dynamical structure and definition of energy in General Relativity. *Physical Review*, 116(5):1322–1330, 1959.

[5] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Canonical variables, expressions for energy, and the criteria for radiation in General Relativity. *Il Nuovo Cimento*, 15(3):487–491, 1960.

[6] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Canonical variables for General Relativity. *Physical Review*, 117(6):1595–1602, 1960.

[7] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Consistency of the canonical reduction of General Relativity. *Journal of Mathematical Physics*, 1(5):434–439, 1960.

[8] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Energy and the criteria for radiation in General Relativity. *Physical Review*, 118(4):1100–1104, 1960.

[9] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Finite self-energy of classical point particles. *Physical Review Letters*, 4(7):375–377, 1960.

[10] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Gravitational-electromagnetic coupling and the classical self-energy problem. *Physical Review*, 120(1):313–320, 1960.

[11] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Interior Schwarzschild solutions and interpretation of source terms. *Physical Review*, 120(1):321–324, 1960.

[12] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Note on positive-definiteness of the energy of the gravitational field. *Annals of Physics*, 11(1):116–121, 1960.

[13] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Coordinate invariance and energy expressions in General Relativity. *Physical Review*, 122(3):997–1006, 1961.

[14] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Heisenberg representation in classical General Relativity. *Il Nuovo Cimento*, 19(4):668–681, 1961.

[15] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Wave zone in General Relativity. *Physical Review*, 121(5):1556–1566, 1961.

[16] Richard Arnowitt, Stanley Deser, and Charles W. Misner. The dynamics of general relativity. In Louis Witten, editor, *Gravitation: an introduction to current research*, pages 227–265. John Wiley & Sons, Inc., New York and London, 1962. Online available as arXiv:gr-qc/0405109.

[17] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Republication of: The Dynamics of General Relativity. *General Relativity and Gravitation*, 40(9):1997–2027, 2008. Republication as *Golden Oldie* with some minor corrections. Online available as arXiv:gr-qc/0405109.

[18] Abhay Ashtekar. New Hamiltonian formulation of general relativity. *Physicsl Review D*, 36(6):1587–1602, 1987.

[19] Abhay Ashtekar, Aiyalam P. Balachandran, and Sang G. Jo. The CP problem in quantum gravity. *International Journal of Modern Physics A*, 4(6):1493–1514, 1989.

[20] Abhay Ashtekar and Anne Magnon-Ashtekar. On conserved quantities in general relativity. *Journal of Mathematical Physics*, 20(5):793–800, 1979.

[21] Ralph F. Baierlein, David H. Sharp, and John A. Wheeler. Three-dimensional geometry as carrier of information about time. *Physical Review*, 126(5):1864–1865, 1962.

[22] Fernando Barbero. Real Ashtekar variables for Lorentzian signature space-times. *Physical Review D*, 51(10):5507–5510, 1995.

[23] Julian B. Barbour. *Absolute or Relative Motion? Volume 1: The Discovery of Dynamics*. Cambridge University Press, Cambridge, 1989.

[24] Julian B. Barbour. The timelessness of quantum gravity: II. the appearance of dynamics in static configurations. *Classical and Quantum Gravity*, 11(12):2875–2896, 1994.

[25] Robert Bartnik. The mass of an asymptotically flat manifold. *Communications on Pure and Applied Mathematics*, 39(5):661–693, 1986.

[26] Robert Bartnik and Gyula Fodor. On the restricted validity of the thin sandwich conjecture. *Physical Review D*, 48(8):3596–3599, 1993.

[27] Thomas W. Baumgarte and Stuart L. Shapiro. *Numerical Relativity. Solving Einstein's Equations on the Computer*. Cambridge University Press, Cambridge, 2010.

[28] Robert Beig. Arnowitt-Deser-Misner energy and g_{00}. *Physics Letters A*, 69(3):153–155, 1978.

[29] Robert Beig and Niall Ó Murchadha. The Poincaré group as symmetry group of canonical general relativity. *Annals of Physics*, 174:463–498, 1987.

[30] Elliot P. Belasco and Hans C. Ohanian. Initial conditions in General Relativity: Lapse and shift formulation. *Journal of Mathematical Physics*, 10(8):1503–1507, 1969.

[31] Lydia Bieri and Nina Zipser. *Extensions of the Stability Theorem of the Minkowski Space in General Relativity*. Ams/Ip Studies in Advanced Mathematics. American Mathematical Society, Providence, Rhode Island, 2009.

[32] Donato Bini, Giampiero Esposito, Claus Kiefer, Manuel Krämer, and Francesco Pessina. On the modification of the cosmic microwave background anisotropy spectrum from canonical quantum gravity. *Physical Review D*, 87(10):104008 (13 pages), 2013.

[33] Christian Blohmann, Marco Cezar Barbosa Fernandes, and Alan Weinstein. Groupoid symmetry and constraints in General Relativity. *Communications in Contemporary Mathematics*, 15(01):1250061 (25 pages), 2013.

[34] Martin Bojowald. *Canonical Gravity and Applications. Cosmology, Black Holes, Quantum Gravity*. Cambridge University Press, Cambridge, 2011.

[35] Martin Bordemann, Hans-Christian Herbig, and Stefan Waldmann. BRST cohomology and phase space reduction in deformation quantisation. *Communications in Mathematical Physics*, 210(1):107–144, 2000.

[36] Jeffrey M. Bowen. General form for the longitudinal momentum of a spherically symmetric source. *General Relativity and Gravitation*, 11(3):227–231, 1979.

[37] Jeffrey M. Bowen and James W. Jr. York. Time-asymmetric initial data for black holes

and black-hole collisions. *Physical Review D*, 21(8):2047–2056, 1980.

[38] Steven Brandt and Bernd Brügmann. A simple construction of initial data for multiple black holes. *Physical Review Letters*, 78(19):3606–3609, 1997.

[39] Yvonne Choquet-Bruhat. *General Relativity and the Einstein Equations*. Oxford University Press, Oxford, 2009.

[40] Demetrios Christodoulou. The chronos principle. *Il Nuovo Cimento*, 26 B(1):67–93, 1975.

[41] Demetrios Christodoulou and Mauro Francavigila. The geometry of the thin-sandwich problem. In Jürgen Ehlers, editor, *Isolated Gravitating Systems in General Relativity*, volume LXVII of *Proceedings of the International School of Physics "Enrico Fermi"*, pages 480–497. North-Holland Publ. Comp., Amsterdam, 1979.

[42] Demetrios Christodoulou and Niall Ó Murchadha. The boost problem in general relativity. *Communications in Mathematical Physics*, 80(2):271–300, 1981.

[43] Piotr Chruściel. A remark on the positive-energy theorem. *Classical and Quantum Gravity*, 3(6):L115–L121, 1986.

[44] Piotr Chruściel. On angular momentum at spatial infinity. *Classical and Quantum Gravity*, 4(6):L205, 1987.

[45] Piotr T. Chruściel and Helmut Friedrich, editors. *The Einstein Equations and the Large Scale Behaviour of Gravitational Fields*. Birkhäuser Verlag, Basel, 2004.

[46] Gerald M. Clemence. On the system of astronomical constants. *The Astronomical Journal*, 53(6):169–179, 1948.

[47] Gerald M. Clemence. Astronomical time. *Reviews of Modern Physics*, 29(1):2–8, 1957.

[48] Gerald M. Clemence. The concept of ephemeris time: A case of inadvertent plagiarism. *Journal for the History of Astronomy*, 2:73–79, 1971.

[49] Justin Corvino and Haotian Hu. On the center of mass of isolated systems. *Classical and Quantum Gravity*, 25(8):085008 (18 pp), 2008.

[50] Sergio Dain. Positive energy theorems in general relativity. In Abhay Ashtekar and Vesselin Petkov, editors, *Springer Handbook of Spacetime*, pages 363–380. Springer Verlag, Berlin, 2014.

[51] V.I. Denisov and V.O. Solov'ev. The energy determined in general relativity on the basis of the traditional Hamiltonian approach does not have physical meaning. *Theoretical and Mathematical Physics*, 56(2):832–841, 1983.

[52] Bryce Seligman DeWitt. Quantum theory of gravity. I. The canonical theory. *Physical Review*, 160(5):1113–1148, 1967. Erratum, ibid. 171(5):1834, 1968.

[53] Paul A. M. Dirac. Generalized Hamiltonian dynamics. *Proceeedings of the Royal Society of London A*, 246(1246):326–332, 1958.

[54] Paul A. M. Dirac. The theory of gravitation in Hamiltonian form. *Proceeedings of the Royal Society of London A*, 246(1246):333–343, 1958.

[55] Paul A.M. Dirac. *Lectures on Quantum Mechanics*. Belfer Graduate School of Science, Monographs Series Number Two. Yeshiva University, New York, 1964.

[56] Albert Einstein and Wolfgang Pauli. On the non-existence of regular stationary solutions of relativistic field equations. *Annals of Mathematics*, 44(2):131–137, 1943.

[57] Arthur E. Fischer and Jerrold E. Marsden. The Einstein equations of evolution – a geometric approach. *Journal of Mathematical Physics*, 13(4):546–568, 1972.

[58] Arthur E. Fischer and Jerrold E. Marsden. The initial value problem and the dynamical formulation of general relativity. In Stephen W. Hawking and Werner Israel, editors, *General Relativity. An Einstein centenary survey*, pages 138–211. Cambridge University Press, Cambridge, 1979.

[59] Arthur E. Fischer and Jerrold E. Marsden. Topics in the dynamics of general relativity. In Jürgen Ehlers, editor, *Isolated Gravitating Systems in General Relativity*, volume LXVII of *Proceedings of the International School of Physics "Enrico Fermi"*, pages 322–395. North-Holland Publ. Comp., Amsterdam, 1979.

[60] Sebastian Fischetti, William Kelly, and Donald Marolf. Conserved charges in asymptotically (locally) AdS spacetimes. In Abhay Ashtekar and Vesselin Petkov, editors, *Springer Handbook of Spacetime*, pages 381–407. Springer Verlag, Berlin, 2014.

[61] Hans Freudenthal. Über die Enden topologischer Räume und Gruppen. *Mathematische Annalen*, 33(1):692–713, 1931.

[62] Simonetta Frittelli. Note on the propagation of the constraints in standard 3+1 general relativity. *Physical Review D*, 55(10):5992–5996, 1997.

[63] Dennis Gannon. Singularities in nonsimply connected space-times. *Journal of Mathematical Physics*, 16(12):2364–2367, 1975.

[64] Robert Geroch. Spinor structure of spacetimes in general relativity. I. *Journal of Mathematical Physics*, 9(11):1739–1744, 1968.

[65] Robert Geroch. Domain of dependence. *Journal of Mathematical Physics*, 11(2):437–449, 1970.

[66] Gary W. Gibbons. The elliptic interpretation of black holes and quantum mechanics. *Nuclear Physics*, B 98:497–508, 1986.

[67] Domenico Giulini. 3-manifolds for relativists. *International Journal of Theoretical Physics*, 33:913–930, 1994.

[68] Domenico Giulini. Ashtekar variables in classical general relativity. In J. Ehlers and H. Friedrich, editors, *Canonical Gravity: From Classical to Quantum*, volume 434 of *Lecture Notes in Physics*, pages 81–112. Springer Verlag, Berlin, 1994.

[69] Domenico Giulini. Asymptotic symmetry groups of long-ranged gauge configurations. *Modern Physics Letters A*, 10(28):2059–2070, 1995.

[70] Domenico Giulini. What is the geometry of superspace? *Physical Review D*, 51(10):5630–5635, 1995.

[71] Domenico Giulini. The generalized thin-sandwich problem and its local solvability. *Journal of Mathematical Physics*, 40(5):2470–2482, 1999.

[72] Domenico Giulini. That strange procedure called quantisation. In D. Giulini, C. Kiefer, and C. Lämmerzahl, editors, *Quantum Gravity: From Theory to Experimental Search*, volume 631 of *Lecture Notes in Physics*, pages 17–40. Springer Verlag, Berlin, 2003.

[73] Domenico Giulini. Mapping-class groups of 3-manifolds in canonical quantum gravity. In Bertfried Fauser, Jürgen Tolksdorf, and Eberhard Zeidler, editors, *Quantum Gravity: Mathematical Models and Experimental Bounds*. Birkhäuser Verlag, Basel, 2007. Online available at ⟨arxiv.org/pdf/gr-qc/0606066⟩.

[74] Domenico Giulini. The superspace of geometrodynamics. *General Relativity and Gravitation*, 41(4):785–815, 2009.

[75] Domenico Giulini. Energy-momentum tensors and motion in special relativity. In D. Pützfeld, C. Lämmerzahl, and B. Schutz, editors, *Equations of Motion in Relativistic*

Gravity, volume 179 of *Fundamental Theories of Physics*. Springer Verlag, Dordrecht, 2015. arXiv: 1502.03930.

[76] Mark Gotay, James Nester, and George Hinds. Presymplectic manifolds and the Dirac-Bergmann theory of constraints. *Journal of Mathematical Physics*, 19(11):2388–2399, 1978.

[77] Éric Gourgoulhon. *3+1 Formalism in General Relativity*, volume 846 of *Lecture Notes in Physics*. Springer Verlag, Berlin, 2012.

[78] Mikhael Gromov and Blaine Lawson. Positive scalar curvature and the Dirac operator on complete Riemannian manifolds. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 58(1):83–196, 1983.

[79] Stephen W. Hawking and George F.R. Ellis. *The Large Scale Structure of Spacetime*. Cambridge University Press, Cambridge, 1973.

[80] Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems*. Princeton University Press, Princeton, New Jersey, 1992.

[81] Sergio A. Hojman, Karel Kuchař, and Claudio Teitelboim. Geometrodynamics regained. *Annals of Physics*, 96:88–135, 1976.

[82] Petr Hořava. Quantum gravity at a Lifshitz point. *Physical Review D*, 79(8):084008 (15 pages), 2009.

[83] Lan-Hsuan Huang. On the center of mass in general relativity. arXiv:1101.0456. Contribution to the Proceedings of Fifth International Congress of Chinese Mathematicians.

[84] Gerhard Huisken and Yau Shing-Tung. Definition of center of mass for isolated physical systems and unique foliations by stable spheres with constant mean curvature. *Inventiones mathematicae*, 124(1-3):281–311, 1996.

[85] Giorgio Immirzi. Real and complex connections for canonical gravity. *Classical and Quantum Gravity*, 14(10):L117–L181, 1997.

[86] James Isenberg. The initial-value problem in general relativity. In Abhay Ashtekar and Vesselin Petkov, editors, *Springer Handbook of Spacetime*, pages 303–321. Springer Verlag, Berlin, 2014.

[87] Christopher J. Isham. *Theta*–states induced by the diffeomorphism group in canonically quantized gravity. In M.J. Duff and C.J. Isham, editors, *Quantum Structure of Space and Time*, Proceedings of the Nuffield Workshop, August 3-21 1981, Imperial College London, pages 37–52. Cambridge University Press, London, 1982.

[88] Christopher J. Isham and Karel V. Kuchař. Representations of spacetime diffeomorphisms. I. Canonical parametrized field theories. *Annals of Physics*, 164:288–315, 1985.

[89] Christopher J. Isham and Karel V. Kuchař. Representations of spacetime diffeomorphisms. II. Canonical geometrodynamics. *Annals of Physics*, 164:316–333, 1985.

[90] Carl Gustav Jacob Jacobi. *Vorlesungen über Dynamik*. Georg Reimer, Berlin, 1866.

[91] Jerry L. Kazdan and Frank W. Warner. Scalar curvature and conformal deformation of Riemannian structure. *Journal of Differential Geometry*, 10(1):113–134, 1975.

[92] Claus Kiefer. *Quantum Gravity*, volume 155 of *International Series of Monographs on Physics*. Clarendon Press, Oxford, third edition, 2012.

[93] Claus Kiefer and Manuel Krämer. Quantum gravitational contributions to the cosmic microwave background anisotropy spectrum. *Physical Review Letters*, 108(2):021301 (4 pages), 2012.

[94] Arthur Komar. Covariant conservation laws in general relativity. *Physicsl Review*, 113(3):934–936, 1959.

[95] Martin D. Kruskal. Maximal extension of Schwarzschild metric. *Physical Review*, 119(5):1743–1745, 1960.

[96] Karel Kuchař. Geometrodynamics regained: A Lagrangian approach. *Journal of Mathematical Physics*, 15(6):708–715, 1974.

[97] André Lichnerowicz. *Théories Relativistes de la Gravitation et de l'Électromagnétisme*. Masson et Cie, Paris, 1955.

[98] David Lovelock. The four-dimensionality of space and the Einstein tensor. *Journal of Mathematical Physics*, 13(6):874–876, 1972.

[99] Fotini G. Markopoulou. Gravitational constraint combinations generate a Lie algebra. *Classical and Quantum Gravity*, 13(9):2577–2584, 1996.

[100] John W. Milnor and John W. Stasheff. *Characteristic Classes*, volume 76 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, New Jersey, 1974.

[101] Charles Misner and John A. Wheeler. Classical physics as geometry: Gravitation, electromagnetism, unquantized charge, and mass as properties of curved empty space. *Annals of Physics*, 2:525–660, 1957.

[102] Niall Ó Murchadha. Total energy momentum in general relativity. *Journal of Mathematical Physics*, 27(8):2111–2128, 1986.

[103] Niall Ó Murchadha and Krzysztof Roszkowski. Embedding spherical spacelike slices in a Schwarzschild solution. *Classical and Quantum Gravity*, 23(2):539–547, 2006.

[104] Jorge Pullin. Editorial note to R. Arnowitt, S. Deser, C. W. Misner: The dynamics of general relativity. *General Relativity and Gravitation*, 40(9):1989–1995, 2008.

[105] Tulio Regge and Claudio Teitelboim. Role of surface integrals in the Hamiltonian formulation of general relativity. *Annals of Physics*, 88:286–318, 1974.

[106] Alan Rendall. *Partial Differential Equations in General Relativity*, volume 16 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, New York, 2008.

[107] Carlo Rovelli. *Quantum Gravity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2004.

[108] Carlo Rovelli. "Forget Time". arXiv: 0903.3832, 2009.

[109] Joseph Samuel. Is Barbero's Hamiltonian formulation a gauge theory of Lorentzian gravity? *Classical and Quantum Gravity*, 17(20):L141–L148, 2000.

[110] Michael Spivak. *A Comprehensive Introduction to Differential Geometry I-V*. Publish or Perish, Inc., Wilmington, Delaware, 1979.

[111] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, Princeton, New Jersey, 1951.

[112] Claudio Teitelboim. How commutators of constraints reflect the spacetime structure. *Annals of Physics*, 79(2):542–557, 1973.

[113] Thomas Thiemann. *Modern Canonical Quantum General Relativity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2007.

[114] Richard C. Tolman. On the use of the energy-momentum principle in General Relativity. *Physical Review*, 35(8):875–895, 1930.

[115] Richard C. Tolman. *Relativity, Thermodynamics, and Cosmology*. Oxford University Press, Oxford, 1934.

[116] John A. Wheeler. Geometrodynamics and the issue of the final state. In Cecile M.

DeWitt and Bryce S. DeWitt, editors, *Relativity, Groups and Topology*, 1963 Les Houches Lectures, pages 317–520. Gordon and Breach, New York, 1964.

[117] John Henry Constantine Whitehead. The immersion of an open 3-manifold in Euclidean 3-space. *Proceedings of the London Mathematical Society (3)*, 11(1):81–90, 1961.

[118] Donald Witt. Vacuum space-times that admit no maximal slices. *Physical Review Letters*, 57(12):1386–1389, 1986.

[119] James W. York. Kinematics and dynamics of general relativity. In Larry Smarr, editor, *Sources of Gravitational Radiation*, pages 83–126. Cambridge University Press, Cambridge, 1979.

[120] James W. York. Conformal "thin-sandwich" data for the initial-value problem of general relativity. *Physical Review Letters*, 82(7):1350–1353, 1999.

Index