Counterfactuals in the initial value formulation of general relativity

José Luis Jaramillo and Vincent Lam

Abstract

How precisely to understand and evaluate counterfactuals can be an intricate issue. The aim of this article is to examine a new set of difficulties for evaluating counterfactuals that arise in the context of the dynamical spacetimes described by the theory of general relativity. The initial value formulation provides us with a methodology to pin down the specific combination of features of the theory at the origin of the difficulties, namely non-linearity and certain non-local aspects (typically captured by ellipticity at the analytical level), in particular when combined with the global and/or quasi-local character of physical quantities in general relativity. Finally, we connect the philosophical question about counterfactuals with concrete applied physical issues in general relativity about extracting meaningful predictions by constructing appropriate initial data sets, leading us to question the very suitability of the Cauchy approach to fully account for the explanatory power of the theory.

1 Introduction

Counterfactuals play an important role in philosophy of science (as well as in science itself), especially in debates about laws of nature and causation. For instance, it is often considered that a crucial aspect of what laws of nature are is their ability to ground counterfactual claims. Similarly, counterfactuals are central to prominent analyses of causation.

How precisely to understand and evaluate counterfactuals can be an intricate issue. The aim in this article is to examine a new set of difficulties for evaluating counterfactuals that arise in the context of the theory of general relativity (GR). The roots of the difficulties lie in the specific dynamical nature of GR spacetime: there is in general no physical (non-dynamical) background structure in GR with respect to which counterfactuals can be evaluated. This general point has been nicely highlighted in (Curiel [2015]). Relying on the fact that the limiting family of Schwarzschild spacetimes parametrized by mass shrinking to zero has no unique limit (see Geroch [1969]), he convincingly argues that counterfactuals involving the removal of some matter in a GR spacetime are therefore ambiguous. So, for instance, counterfactuals of the form 'if the sun would disappear, then the planets would behave in such and such ways' are difficult to rigorously evaluate in the GR context. Curiel's important considerations only aim 'to draw attention to this serious problem', and so remain at a rather general level.

We aim to go a step further by investigating more precisely the difficulties to evaluate counterfactuals in the initial value formulation of general relativity. Beside the fact that the initial value formulation is in many ways the natural setting for evaluating counterfactuals, it allows us to connect with actual concerns of relativists working on astrophysical problems (for example, in numerical relativity approaches) as well as with foundational work in (mathematical) GR. This will enable us to pin down more precisely the combination of specific features (non-linearity and ellipticity) creating the difficulties,[1] and also to discuss new perspectives and new tools to address them.

2 Maudlin's Recipe

Understanding the antecedent of a counterfactual in terms of appropriate modifications in some relevant initial data is a very natural move (to a certain degree, it is intuitively encoded in the standard scientific as well as everyday practice for evaluating scientific and everyday counterfactual claims). Maudlin ([2007], ch. 1) provides a good example of a systematic strategy for evaluating (and grounding) counterfactuals that crucially relies on appropriate modifications in some relevant initial data. Here is an example of a counterfactual claim that Maudlin examines: 'If the bomb dropped on Hiroshima had contained titanium instead of uranium it would not have exploded' (Maudlin [2007], p. 22). In order to ground the truth value of this counterfactual (using appropriate laws of nature), Maudlin ([2007], pp. 22-3) proposes a 'three-step recipe' that makes heavy use of the initial value formulation:[2]

- Step 1: 'choose a Cauchy surface that cuts through the actual world and that intersects the bomb about the time it was released from the plane. All physical magnitudes take some value on this surface.'
- Step 2: 'construct a Cauchy surface just like the one in Step 1 save that the physical magnitudes are changed in this way: uranium is replaced with titanium in the bomb.'
- Step 3: 'allow the laws to operate on this Cauchy surface with the new boundary values generating a new model. In that model, the bomb does not explode. Ergo (if we have got the laws right, etc.) the counterfactual is true.'

How this recipe proceeds for evaluating a given counterfactual is rather straightforward: the central idea is to translate the antecedent of the counterfactual into appropriate changes in relevant Cauchy data (the second step is therefore central). At this point, it is important to highlight two crucial ingredients in Maudlin's recipe, both needed in the second step.

[1] In a certain sense, the structures introduced by the initial value formulation allow to identify more precisely how the difficulties manifest.

[2] Besides the fact that evaluating counterfactuals in terms of changes in initial data is quite natural and intuitive in its own right, it is particularly suited for Maudlin, since he takes laws of nature—and in particular what he calls 'fundamental laws of temporal evolution'—to be ontologically primitive. See also (Hall [2015]) for the link between counterfactuals and initial data (in a Humean context).

First, the Cauchy data have to be modified in some prescribed unique way—that is, given some appropriate Cauchy surface, the antecedent of the counterfactual has to be translated into initial data changes in some a priori specified unique way. This ingredient aims to avoid unwanted ambiguities that could make the whole recipe fail (as in the case where different modifications corresponding to the same antecedent would lead to different verdicts about the truth value of the counterfactual).

Second, only the relevant Cauchy data should be modified—the rest should be left unaltered. As Maudlin ([2007], p. 24) puts it, this is a 'ceteris paribus condition: leave everything else the same'. The idea is clear enough: modifying irrelevant Cauchy data (with respect to the antecedent) might interfere with the counterfactual evaluation in a way that makes the latter untrustworthy and even meaningless.

These two crucial ingredients for evaluating counterfactuals may be lacking simply because the antecedent of the counterfactual under consideration is too vaguely specified (to some extent, the philosophy literature on counterfactuals sometimes contains discussion of contrived and sometimes too vaguely specified examples). The point we want to stress in this article is that, even if the antecedent is specified precisely enough (as in Maudlin's example above or as in standard scientific counterfactual claims[3]) we may lack these two crucial ingredients in GR because of certain features related to the dynamical nature of general relativistic spacetime, specifically non-linearity and global/quasi-local aspects associated with diffeomorphism invariance. Indeed, the latter entails the presence of geometrical constraints, displaying ellipticity at the analytical level[4]. We aim to articulate how this dynamical nature manifests itself concretely through the lack of these two ingredients in the context of the initial value formulation of the theory.

3 The Initial Value Problem Of General Relativity

In many ways, the initial value problem is about the specification of appropriate initial data for GR. This is a tricky issue since GR is not explicitly a theory about the temporal evolution of some dynamical system. Rather, the Einstein field equations relate the geometry of spacetime (encoding the gravitational and inertial structure) with its (non-gravitational) matter 'content'. However, certain spacetimes—in particular globally hyperbolic spacetimes (on which we focus here)[5]—can be considered as resulting from the evolution of some 3-dimensional space (for instance, corresponding to the universe at some time). The general idea is to cast the Einstein equations within some spatial foliation of spacetime as evolution equations for some set of initial data (for instance, encoding the state of the universe at some 'initial' time).

Considering for simplicity the purely gravitational (vacuum) sector of the theory, and given a spatial foliation of a spacetime (M, g), the 3-metric γ_{ij} and the extrinsic curvature K_{ij} on a leaf Σ of the foliation constitute an appropriate choice of initial data.[6] A crucial feature is then that not all Einstein equations correspond to evolution equations: this can be understood in relation to

[3] In the context of GR, examples of standard counterfactuals include conditional claims of the following type: if the black holes in the astrophysical compact binary corresponding to the GW150914 detection had been more massive, the frequencies observed by Advanced LIGO would have been lower.

[4] In our discussion we will focus on ellipticity for methodological reasons, since it captures the global relevant aspects (see footnote 13) in a form particularly well suited for the present analysis and discussion.

[5] The initial value formulation can be extended to certain non-globally hyperbolic situations (Friedman [2004]). Here we aim to consider issues arising for the evaluation of counterfactuals already in the friendliest (that is, globally hyperbolic) GR spacetimes in a Cauchy initial data setting.

[6] For the sake of readability, and in order to focus on the conceptual aspects related to counterfactuals, we privilege in this paper a qualitative discussion over technical details; for the latter, we refer the interested reader to (Isenberg [2014])—which we broadly follow for the technical aspects in this article—as well as references therein.

the (diffeomorphism) gauge invariance of the theory, namely cast in terms of the (Gauss-Codazzi) consistency conditions guaranteeing that the 3-slices in the foliation actually build up the spacetime 4-manifold. Indeed, given a spatial foliation, Einstein equations decompose into six evolution equations and four constraint equations for the initial data. In this respect, the situation is very similar to electrodynamics where the Maxwell equations also separate into (the Gauss) constraint and evolution equations.

So, from the Einstein equations, one obtains the following four (vacuum) constraints for the initial data (γ_{ij}, K_{ij}):

$$R + K^2 - K_{ij}K^{ij} = 0 \quad (1)$$

$$\nabla_j(K^{ij} - \gamma^{ij}K) = 0 \;, \quad (2)$$

where $K = K_i{}^i$, ∇ is the covariant derivative associated with γ_{ij} and R is the corresponding Ricci scalar. Equation (1) is called the Hamiltonian constraint and (2) is called the momentum constraint. The important point is that the initial data (γ_{ij}, K_{ij}) have to satisfy these constraints—and so cannot be freely prescribed—if their evolution is to build up a spacetime consistent with the Einstein equations. That a unique (up to diffeomorphism) solution to the Einstein equations exists for any smooth set (γ_{ij}, K_{ij}) satisfying the constraints (1)-(2) is guaranteed by a powerful ('well-posedness') theorem due to Choquet-Bruhat and Geroch ([1969]).

This fundamental well-posedness result does not say anything about how to specify the initial data; but such specification is clearly crucial for making predictions, and, more specifically, for evaluating counterfactual claims. Indeed, the second step in Maudlin's recipe (section 2) involves modifying the initial data in some appropriate and precise way corresponding to the antecedent of the counterfactual to be evaluated (for instance, 'uranium is replaced with titanium in the bomb'), and such that the constraints are still satisfied. As we will see, the nature of these latter constraints renders this task difficult in GR.

4 Conformal Strategies For Solving The Constraints

A first issue when specifying appropriate initial data comes from the fact that the system of constraint equations (1)-(2) for (γ_{ij}, K_{ij}) is underdetermined; there is some freedom in choosing which components of γ_{ij} and K_{ij} are constrained and which ones are freely specifiable. We start by discussing a standard strategy for solving the constraints that makes use of conformal techniques. The main idea is to decompose the initial data into (conformal) 'free data' and 'determined data', such that the geometric constraint equations are written as an elliptic partial differential equation system for the determined data. In this reduction process, the original geometric problem is turned into one of analytic character.[7]

For the sake of concreteness, we consider the York-Lichnerowicz decomposition that involves a conformal decomposition of the 3-metric γ_{ij} into a conformal factor ϕ and a conformal 3-metric $\tilde{\gamma}_{ij}$, $\gamma_{ij} = \phi^4\tilde{\gamma}_{ij}$, together with a conformal 'transverse-traceless' decomposition of the extrinsic curvature K_{ij}. The set of partial differential equations that arise from the York-Lichnerowicz decomposition is a challenging non-linear elliptic system, for which known existence and uniqueness results are limited to specific classes. An existence and uniqueness result lacks in the general case, in particular applying to the actual generic situations encountered in astrophysical and cosmological scenarios.

[7]To avoid confusion with the common philosophical understanding: 'reduction' is the technical term used in the relevant literature here and only refers to this geometric-analytic transition.

The system simplifies (in the considered vacuum case) if a constant mean curvature gauge condition for the initial data is adopted, that is, $\nabla_i K = 0$.[8] The momentum constraint (2) decouples from the Hamiltonian constraint (1) and can be solved independently of it. Introducing the obtained solution into (1) then reduces the Hamiltonian constraint to the Lichnerowicz equation for the conformal factor ϕ:

$$\tilde{\Delta}\phi - \frac{1}{8}\tilde{R}\phi + \frac{1}{8}\tilde{A}^{ij}\tilde{A}_{ij}\phi^{-7} + \frac{1}{12}K^2\phi^5 = 0 \ , \tag{3}$$

where the Laplacian $\tilde{\Delta}$ and the Ricci scalar $\tilde{R}$ are associated with the conformal 3-metric $\tilde{\gamma}_{ij}$, and where $\tilde{A}_{ij}$ is the conformal tracefree extrinsic curvature (we have the decomposition $K_{ij} = \phi^{-2}\tilde{A}_{ij} + \frac{1}{3}K\gamma_{ij}$). In this process the conformal 3-metric $\tilde{\gamma}_{ij}$ and a part of the conformal tracefree extrinsic curvature $\tilde{A}_{ij}$ (namely the so-called transverse and traceless part $\tilde{A}^{TT}_{ij}$) can be freely specified (free data), whereas the conformal factor ϕ is the determined data, solution of the equation (3).[9]

For our purpose in this paper, it is crucial to highlight two very important features of the Lichnerowicz equation: it is an elliptic and non-linear partial differential equation. These features are also present in the case without the constant mean curvature gauge condition. More generally, all existing approaches to the problem that have proven to be effective in the construction of initial data in generic scenarios include elliptic and non-linear features.[10] It is in particular the case, on the one hand, of the family of conformal decomposition approaches allowing for the ab initio construction of initial data, such as the one described above or the related conformal thin sandwich decomposition (York [1999]; Pfeiffer and York [2003]) and, on the other hand, of the gluing techniques for combining two previously known solutions to the contraints (see section 7) (Chruściel *et al.* [2005]; Corvino and Schoen [2006]). As we will now discuss, the combination of these non-linear and elliptic properties (the latter standing as the analytical signature of the geometric constraints encoding the here discussed non-local features of the theory) characterizes the heart of the difficulties for evaluating counterfactuals in GR.

[8]Nothing substantial for our discussion hinges on this assumption, which, on the other hand, actually amounts to a gauge choice in the slicing of the 3+1 decomposition; it allows to ease the discussion of the partial differential equation we end up with. Moreover, several very interesting results have been proven using such constant mean curvature data, in particular on closed manifolds (using Yamabe theorem, which ensures that any Riemannian metric on a closed manifold with dimension no less than 3 can be conformally transformed into a metric with constant scalar curvature); for non-constant mean curvature data, much less is known about existence and uniqueness of solutions (see (Isenberg [2014], §16.4) and references therein). As a consequence, dropping the constant mean curvature assumption actually further strengthens the difficulties for evaluating counterfactuals, since then one does not even know whether initial data corresponding to the antecedent of the counterfactual to be evaluated exist (that is, whether appropriate solutions to the constraints exist).

[9]A vector X^i, also subject to an elliptic equation (see table 1) and necessary to reconstruct $\tilde{A}_{ij}$ from $\tilde{A}^{TT}_{ij}$, completes the determined data in the constraints. The discussion is more transparent, without loss of generality, if we focus on Lichnerowicz equation (3).

[10]The underdetermined nature of the constraints allows for different kinds of reductions to a partial differential equation system. Elliptic reductions constitute the most studied and best understood class of reductions, but they are not the only possibility. Parabolic reductions of the Hamiltonian constraint have been indeed studied (Bartnik and Isenberg [2004]) and, more recently, the reduction of the constraints to hyperbolic-algebraic and/or hyperbolic-parabolic systems constitute an active and promising field of research (Rácz [2016]). All these approaches contain an essential 'non-local' aspect, which is the main feature we aim at highlighting here. Since the discussion of this 'non-local' element is (technically) better controlled in elliptic formulations, we adopt the methodological choice of focusing on the well-studied elliptic systems.

5 Non-linearity And Ellipticity At The Heart Of The Difficulties

Let us briefly recapitulate where we are. In order to ground the truth value of counterfactual claims in a general relativistic context following the intuitive procedure captured in Maudlin's recipe (section 2), we have considered the (Cauchy) initial value formulation of GR, within which the Einstein equations give rise to four constraints on the initial data alongside the six evolution equations (section 3). In other terms, initial data have to solve these constraints in order to be consistent with Einstein equations; these constraints are underdetermined and in particular the choice of their unconstrained (free) part determines the physical content of the initial data. However, in their geometric form, the constraints are not well suited for their systematic resolution in the generic case (and more so for concrete applications). The standard way to deal with this situation is to recast the geometric equations as an analytic problem, the elliptic reductions (and notably the conformal decompositions) representing the most systematically developed strategy. Initial data are then constructed by solving a non-linear elliptic partial differential equation system for the determined data (section 4).

Now, given a set of appropriate initial data (that is, solving the constraints) corresponding to some actual situation (as in step 1 of Maudlin's recipe), the second step of Maudlin's recipe amounts to modify the initial data (to find a new solution to the constraint system) according to the antecedent of the counterfactual to be evaluated. In order to do so, we have to translate the antecedent in terms of some prescription for the free data (for instance, the conformal 3-metric $\tilde{\gamma}_{ij}$), then solve the partial differential equation system (for instance, the Lichnerowicz equation)—assuming a solution exists and is unique, which is far from clear in general—for the determined data (for instance, the conformal factor ϕ), and all this such that the resulting reconstructed physical initial data (for instance, γ_{ij} and K_{ij})[11] correspond to the antecedent of the counterfactual.

At this stage, we already see that there are quite a lot of difficulties to overcome for step 2 of Maudlin's recipe to proceed: solving the relevant partial differential equation system can be very tricky indeed (some systems may even have no solution). Here, we want to focus more specifically on fundamental features—ellipticity and non-linearity—of the partial differential equation system that raise in-principle difficulties for evaluating counterfactuals.

In order to be meaningful, we have seen that the second step of Maudlin's recipe requires two crucial ingredients (section 2): first, the initial data have to be modified in some prescribed unique way, and second, only relevant data should be modified, leaving the rest untouched (ceteris paribus condition). We have already seen above that uniqueness of solutions—as their very existence!—may not be guaranteed in general, non-linearity playing a fundamental role in these obstructions. Moreover, such non-linearity of the theory—of the relevant equations—also plays a major role in the difficulties to modify initial data according to some prescription (given by the antecedent of the considered counterfactual). For instance, because of non-linearity, one cannot build a new solution by adding two solutions; in particular, one cannot simply add 'perturbations' to known solutions (in a general relativistic context, the replacing of uranium with titanium in Maudlin's example cannot simply be considered as a perturbation of the actual initial data).[12]

[11] Of course, the initial data should include non-gravitational (matter) initial data in the non-vacuum case; this makes the constraint equations more complicated and adds further difficulties for solving the equations.

[12] A partial ease to this situation is provided by gluing (see section 7), since the latter indeed provides an approach to the construction of a new solution from two previously known ones. However, such ease is only partial. Firstly, because nothing is said of how these two previous sets of data are constructed, a point critical in Maudlin's step 2, where the construction of new data is subject to the specific prescription corresponding to the antecedent of the counterfactual to be evaluated. And secondly, because the gluing approach involves an elliptic step that alters the neighbourhood of the glued data (a 'quasi-local' behaviour in the sense discussed in footnote 13), therefore also affecting the ceteris paribus condition.

Elliptic equations for certain data—in contrast to parabolic or hyperbolic equations—do not describe any evolution for these data; rather they encode global constraints for these data. More generally, elliptic systems in physics typically appear as such constraints on evolutionary systems or as describing equilibrium states, Poisson and Laplace equations being the paradigmatic examples; in general relativity they can also appear as gauge conditions on the evolution (Dain [2006]). Elliptic equations possess a 'non-local' character in the sense that the value of the solutions at a point potentially depend on all prescribed data (such non-locality can be more finely qualified in terms of global and quasi-local features, which turn out to be crucial in GR; see the footnote for details)[13], so that the solution construction involves the resolution of the equations on the full 3-slice as a whole (that is, all the data 'at the same time')—possibly including global boundary and/or asymptotic conditions. This 'non-local' character is shared with parabolic evolutions but is in stark contrast with the locally finite speed propagation in hyperbolic equations, where finite time evolution at a point may concern only a certain part of the data while being 'unaffected' by sufficiently distant data. More specifically, given a spatially localized source of an elliptic system, the resulting solution typically presents a global spatial structure (or at least a non-fully localized one, in some special cases). An example of this in electrostatics is the Coulomb field solution of the Poisson equation, which can be produced by a pointlike charge source but extends to the whole space.

Adding the non-linear ingredient to this fundamental ellipticity of the partial differential equation system that results from the constraints implies loosing the full detailed control on the non-local modifications of the initial data, thus preventing the ceteris paribus condition in Maudlin's recipe to be satisfied. Indeed, in non-linear elliptic equation systems, a neat separation of the global impact on the solution of individual (in particular local) distinct modifications of the free data is in general not possible.[14] Hindrances come in degrees here, from the global long distance impact in the conformal approaches to the quasi-local deformations in gluing (see footnote 12). Such non-local aspects stand as a common feature underlying all approaches to the resolution of the constraints.[15] This can be illustrated with the elliptic and non-linear Lichnerowicz equation (3) (which assumes constant mean curvature for simplicity), where the antecedent of a given counterfactual ('uranium is replaced with titanium in the bomb')[16] cannot be translated in terms of the freely specifiable data $\tilde{\gamma}_{ij}$ without affecting the whole physical 3-metric γ_{ij} via the conformal factor (and constrained data) ϕ, solution to (3).

Therefore, it is actually the combination of non-linearity and non-locality (dressed at the analytical level in the form of ellipticity) in the theory of general relativity that creates the difficulties for constructing appropriate initial data corresponding to some prescription.

A priori one does not fully control the detailed relation of the constructed solution to the employed building blocks (either free data in the conformal approaches or previously known solutions in gluing), in particular due to the discussed non-local and non-linear properties of the solution.

[13] More precisely, given a modification of the equation (for instance, a modification of the free data) in a compact set $\mathcal{S}$ of the 3-slice, we will refer to the behaviour of the solution as i) global if the latter is modified on the whole slice and as ii) quasi-local if there exists an open neighbourhood $\mathcal{O}$ of compact closure such that the region where the solution is modified is included in $\mathcal{O}$ and the intersection with $\mathcal{O} \setminus \mathcal{S}$ is non-empty. In other words, a global feature of the solution involves the whole slice whereas a quasi-local one only affects a bounded region but strictly extending the modified zone. Note that the quasi-local notion here is in line with the standard use in GR (Szabados [2009]); see also footnote 23.

[14] Thanks to an anonymous referee for pressing us on this point.

[15] Likewise, the parabolic and hyperbolic reductions in footnote 10 share this global/quasi-local character, though cast in different specific forms. Essentially, this feature stems from the fact that the Ricci scalar R of the spatial slice in the Hamiltonian constraint (1) corresponds to a second order differential operator whose principal part is given in terms of a Riemann metric, namely γ_{ij}.

[16] Maudlin's example requires to consider non-vacuum GR; as already mentioned, this only makes things worse.

This can be nicely illustrated by considering theories displaying one of these features but not the other: in such theories, the above mentioned difficulties do not arise. We have illustrated the situation in table 1, with the aid of three historically and logically 'ordered' gravitational theories: i) Newton(-Cartan) theory, elliptic but (essentially) linear, ii) Nordström (second) theory of gravitation, non-linear but without elliptic component, and iii) GR, both non-linear and with an elliptic sector. These gravitational examples are complemented with Maxwell's theory of electromagnetism, with elliptic component but linear, since its somewhat intermediate structure between Newton-Cartan and GR permits to stress the relevant points.[17]

More specifically, in Newton-Cartan theory the field equations do not describe an evolving system and can actually be cast as an elliptic system at the analytical level. In particular, the theory can be expressed in terms of an antisymmetric field (closed $2-$form $\kappa_{\mu\nu}$), whose 'electric' part κ_{0i} leads to an elliptic linear equation for a scalar field potential ϕ at a given time slice (see Rueede and Straumann [1997]). This is indeed the standard (Newton) gravitational Poisson equation slightly generalized through a repulsive gravity term in the source, namely a quadratic term in the magnetic part κ_{ij}.[18] The linearity of the Poisson equation permits a full control of the analytic properties of the solutions, in particular the separation of the effects of different modifications. On the other side, very roughly, Nordström (second) theory of gravitation describes a non-constrained propagating scalar field ϕ, with $\phi\Box\phi = -\frac{4\pi G}{c^4}T_\mu{}^\mu$ (where $\Box$ is the standard d'Alembertian).[19] Although the evolution is non-linear, the absence of constraints allows the essentially free prescription of initial data. As a consequence, in neither of these gravitational theories does one face the GR difficulties for constructing prescribed Cauchy data, in the former case because of the absence of non-linearity and in the latter case because of the absence of constraints and the associated ellipticity. Therefore, in both Newton-Cartan and Nordström theory we can fully implement Maudlin's step 2. An interesting intermediate case in the path towards GR is provided by Maxwell's theory of electromagnetism. Such theory describes a system in evolution, but, in contrast with Nordström theory, it possesses an elliptic sector due to the Gauss constraint, namely the Poisson equation with the charge density as a source. However, as in the Newton-Cartan case, the linearity of the theory permits to fully control the construction and manipulation of initial data. As in the previous gravitational examples, no problem therefore arises in the implementation of Maudlin's step 2.[20] As we have stressed, the situation changes dramatically in GR due to the simultaneous presence of non-linearity and ellipticity.

We have illustrated this discussion in table 1 (note that ϕ stands for the gravitational (scalar) potential in Newton-Cartan theory, a scalar field in Nordström theory, the electric potential in Maxwell's theory and a conformal factor in general relativity).[21]

---

[17]We are grateful to an anonymous referee for suggesting that we also consider Newton-Cartan theory.

[18]Such magnetic term κ_{ij} can be written in terms of a 'dual' vector potential A^i satisfying a decoupled elliptic equation, thus entering in the quadratic source of the Poisson equation as a fixed term. Therefore the non-linearity of the total elliptic system is mild (we have noted this minor subtlety with an asterisk (*) in the 'Non-linearity' line in table 1). In addition, such a magnetic term is fully absent in isolated systems with appropriate boundary conditions at infinity (see references in Rueede and Straumann [1997]).

[19]See (Norton [1992]) for a fascinating conceptual and historical discussion of Nordström theories of gravitation.

[20]A difference shows up though between, on the one side, Nordström and, on the other side, Newton-Cartan and Maxwell theories. A global aspect in the initial data is actually present in the latter, due to their elliptic components, that is completely absent in the former. However, such globality is fundamentally milder in Newton-Cartan and Maxwell theories than in the non-linear GR case, since the critical role of linearity permits an a priori control of the behaviour of solutions. In particular, they allow for a full control of the construction of physically relevant new solutions, by simply using linear combinations of previous valid data.

[21]The notation in the table is fairly standard; in particular, $C_{\mu\nu\rho\sigma}$ denotes the Weyl tensor and $\tilde{\Delta}_{\mathbb{L}}$ a vectorial Laplacian-like differential operator associated with $\tilde{\gamma}_{ij}$, obtained from the divergence of the 'longitudinal' part of the conformal tracefree extrinsic curvature (see Pfeiffer and York [2003]; details are not relevant for our main purpose).

	Newton-Cartan Gravity	Nordström Gravity	Electromagnetism	General Relativity
Field equations (geometric form)	Galileo connection (closed 2-form $\kappa_{\mu\nu}$): $R_{\mu\nu} = 8\pi G \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\right)$ $\nabla_\mu T^{\mu\nu} = 0$	Propagating scalar field ϕ: $R = \frac{24\pi G}{c^4} T_\mu{}^\mu$ $C_{\mu\nu\rho\sigma} = 0$	Maxwell equations (linear spin-1 field): $\nabla_\mu F^{\mu\nu} = \mu_0 J^\nu$ $\nabla_\mu F_{\nu\rho} + \nabla_\rho F_{\mu\nu} + \nabla_\nu F_{\rho\mu} = 0$	Einstein equations (non-linear spin-2 field): $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$ $\nabla_\mu G^{\mu\nu} = 0$
Hyperbolicity	**No**	**Yes** $\phi \Box \phi = -\frac{4\pi G}{c^4} T_\mu{}^\mu$	**Yes** $\Box A^\mu - R^\mu{}_\nu A^\nu = \mu_0 J^\mu$	**Yes** $\Box h^{\mu\nu} = \text{Non-Lin}^{\mu\nu}(h^{\rho\sigma})$
Ellipticity	**Yes** Field equations in Galilei coordinates: $2\nabla_i {\kappa_0}^i - \kappa_{ij}\kappa^{ij} = 4\pi G\rho$ $\nabla_j {\kappa_i}^j = 0$	**No**	**Yes** Gauss Constraint: $\nabla_i E^i = \frac{\rho_c}{\epsilon_0}$	**Yes** Hamiltonian & Momentum constraints: $R + K^2 - K_{ij}K^{ij} = \frac{16\pi G}{c^2}\rho$ $\nabla_j(K^{ij} - \gamma^{ij} H) = \frac{8\pi G}{c^3} j^i$
	$\Delta\phi = 4\pi G\rho - \kappa_{ij}\kappa^{ij}$ $-\Delta A^i + \nabla^i \nabla_j A^j = 0$		$\Delta\phi = -\frac{\rho_c}{\epsilon_0}$	$\tilde{\Delta}\phi - \frac{1}{8}\tilde{R}\phi = \text{Non-Lin}(\phi)$ $\tilde{\Delta}_{\mathbb{L}} X^i = \text{Non-Lin}^i(\phi)$
Non-linearity	**No***	**Yes**	**No**	**Yes**

Table 1: This table presents the status of the four classical field theories discussed in the text with respect to non-linearity and ellipticity. The absence of one of the two ingredients removes the difficulties to implement prescribed modifications of the initial data (as in the case of Nordström gravity) or puts them under full control (as in the case of the linear Newton-Cartan and Maxwell theories). On the contrary, the simultaneous presence of both elements occurring in GR inevitably introduces fundamental issues in the construction and (prescribed) modification of initial data, specifically concerning existence and uniqueness of solutions to the constraints, as well as sufficient control of non-local modifications of these solutions.

6 Relativistic Infection And Possible Worlds

Within the framework of stochastic theories, Maudlin ([2007], p. 30) actually considers the kind of difficulties that we have highlighted in the previous section and that may prevent any meaningful application of the second step of his initial value recipe for evaluating counterfactuals (because the uniqueness and ceteris paribus ingredients may fail to obtain):

> ‘Let us call any physical magnitude which is unchanged when we apply Step 2 uninfected; those which are changed are, in both the original and new data sets, infected. This distinction may be quite clear in some theories, in others (notably in certain quantum states) the physical magnitude may not be localized and may be so entangled that the infection cannot be localized. In such cases, our intuitions break down.’

In situations where quantum entanglement plays an important role, the ceteris paribus condition may not be satisfied. This is an aspect of a set of well-known difficulties for the analysis of causation and counterfactuals in the quantum context (in particular linked to stochasticity and quantum non-locality).[22] The above discussion demonstrates that this is not only a quantum issue (or an issue about stochasticity), but also a general relativistic one. Namely, as we have discussed, the initial value problem in GR displays global/quasi-local aspects related to the construction of appropriate initial data; moreover, beyond the initial data problem, GR as such presents additional fundamentally non-local issues, which impact the analysis here. More specifically in the present context, the global/quasi-local aspects in the initial data subtly combine with the absence of local densities for the construction of physical quantities. Indeed, as a consequence of such lack of local

[22]See for instance Healey [2009] and Bigaj [2006], as well as references therein.

densities, in generic situations physical quantities have a fundamentally global or at least quasi-local nature in GR.[23] This impacts in a critical way the assessment of Maudlin's step 2 in GR, since the inescapable non-local modifications in the initial data (even if, as it happens in the gluing approach, they do not extend to the full slice) unavoidably 'infect' quasi-local physical magnitudes as well as, of course, global ones.[24] In a slogan form and using Maudlin's characterization: 'there's not only quantum, but also (general) relativistic infection'.

The GR context does not only raise difficulties for Maudlin's recipe, but also for other accounts of counterfactuals, such as Lewis'—one of the most prominent analyses of counterfactuals (Lewis [1973], [1979]).[25] Lewis' account famously relies on a (context-dependent) metric of similarity between worlds: glossing over many subtleties, his strategy for evaluating counterfactuals is to consider the closest possible worlds where the antecedent is true and check the claim there. Lewis ([1979], p. 472) proposes one procedure for measuring similarity between worlds:

(1) 'It is of the first importance to avoid big, widespread, diverse violations of law.

(2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.

(3) It is of the third importance to avoid even small, localized, simple violations of law.

(4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.'

The general idea is pretty clear: a world where there is no big, widespread, diverse violation of laws is more similar (and hence closer) to the actual one than a world where there is such big violation. Furthermore, given two worlds satisfying (1), the one with the biggest spacetime region of perfect match of particular fact with respect to the actual one will be more similar (and hence closer) to the actual world—and so on. This procedure—and Lewis' similarity relation in general—faces many issues (in particular, as already mentioned, in relation to stochasticity and quantum non-locality, see Bigaj [2006]), the details of which need not concern us here.

From the discussion in the previous section, we see that GR brings new, additional difficulties in the concrete implementation of Lewis' procedure; even if this latter does not explicitly rely on the initial value framework, in many concrete implementations, it actually does.[26] Very roughly, the problem is that, in a GR setting, we don't know how to evaluate how close to the actual world is a possible world where the antecedent of the counterfactual under consideration holds (in a sense, and because of the difficulties we have highlighted in this article, these worlds all lie 'far away'). Indeed, as we have seen, a general prescribed modification of small particular facts (following the

[23]Such absence of local densities is rooted in the equivalence principle and the lack of rigid background structures; this has led to the introduction in GR of quasi-local quantities "associated with open subsets of spacetime, whose closure is compact" (Szabados [2009], p. 13). Note the strong analogy between this standard GR notion of quasi-local quantities and the quasi-local behaviour of solutions we have introduced in footnote 13 (hence the same terminology); the intertwining of these two closely related aspects lies at the heart of the difficulties in constructing initial data corresponding to a priori prescribed physical properties.

[24]By contrast, this point is fundamentally absent in Maxwell's theory, where local densities of the field energy do exist.

[25]Despite providing somewhat related analyses of counterfactuals (at least in spirit), Maudlin and Lewis defend radically different metaphysical conceptions, in particular about laws: Maudlin is a primitivist about laws, whereas Lewis is the arch-Humean, considering that laws are reducible to certain regularities or patterns in the total spacetime distribution of (local) particular categorical facts (the 'Humean mosaic').

[26]See Hall [2015] for a recent, updated version of Lewis' account of laws and counterfactuals that explicitly relies on the initial value formulation.

antecedent) may lead to the relevant constraints not to be satisfied, and so to the Einstein equations to be violated (so that (1) in Lewis' procedure would not be satisfied in the first place; these possible worlds would then be considered very 'far away' from the actual one). On the other hand, if we try to guarantee that the antecedent holds such that the constraints are satisfied (not necessarily a tractable task in the generic case!), then the region of perfect match with the actual world may actually be critically affected (including in particular the evaluation of key physical quantities): as we have seen, it is very hard to control prescribed modifications in the Cauchy data, and this affects inevitably the associated quasi-local physical quantities (see footnote 23). As a consequence, the corresponding possible worlds would then also count as being 'far away' from the actual one. From our discussion above, these difficulties that Lewis' similarity relation face should not be surprising since this latter plays an analogous role to the ceteris paribus ingredient in Maudlin's recipe—as noticed by Maudlin ([2007], p. 33) himself.

7 Perspectives

It is important to underline the fact that the difficulties we have discussed in this paper do not only concern the philosophical debate about counterfactuals, but also constitute a crucial concrete problem for extracting useful predictions from GR, as well as for investigating foundational issues in GR. An important example of this is provided by spacetime constructions in numerical relativity performed by solving the GR Cauchy initial data problem, either within the framework of foundational studies about spacetime dynamics (for example, the genericity assessment of Penrose's conformally compactified picture) or in the context of making predictions associated with astrophysical or cosmological phenomena. In the latter case, difficulties are often dealt with by using approximation schemes and/or by taking an 'effective description' point of view. This is for instance the case in the successful prediction of gravitational waveforms from black hole mergers. But even more crucial for the present discussion, it is also the case in the consistency check provided by the comparison of the actually observed signals with the evolution of the appropriate initial data, that is, those 'prescribed' by the specific parameters extracted from the observations.[27] However, the difficulties highlighted in previous sections—for instance, having to do with the particular fundamental dynamical nature of spacetime—suggest that such strategies may not always work, in particular in scenarios where the (non-linear) stability of the system has not been yet established.

Constructing appropriate initial data sets is therefore an area of active current research in GR. Besides the conformal methods we have focused on above, we have also mentioned the so-called gluing procedures. The latter have been developed in order to construct new solutions by combining known solutions (which is precisely what cannot be achieved by simple, straightforward means in conformal approaches because of non-linearity, see section 5). For instance, the Corvino-Schoen asymptotic exterior gluing technique 'allows one to smoothly glue any interior region of an asymptotically Euclidean solution to an exterior region of a slice of a Kerr solution' (Isenberg [2014], p. 317).[28] Moreover, by combining various gluing techniques, Chruściel *et al.* ([2011]) were able to show that 'for any chosen set of N asymptotically Euclidean solutions of the constrains representing black holes, stars, or other astrophysical objects of interest, one can construct a new asymptotically Euclidean solution which includes interior regions of these N chosen solutions, placed as desired

[27] In particular, the characterization of the individual black hole masses and (more dramatically) of the individual spins is plagued with effective prescriptions. In spite of that, the consistency between the reconstructed waveform and the observed one is remarkable. This points indeed towards the need of an in-depth understanding of this often manifested 'effective insensitivity' of the GR dynamics to (sometimes crude) approximations.

[28] This is a major breakthrough with deep foundational meaning, since it grounds the robustness of the conformally compactified picture (namely showing the existence of generic solutions with a smooth null infinity).

(so long as the distances between the bodies are sufficiently large) and with the desired relative momenta' (Isenberg [2014], p. 318).

Of course, these results have their limitations. On the one hand, the distances involved in the N-body problem must be sufficiently large, whereas the Corvino-Schoen technique involves an unknown transition zone; in both cases, the elliptic tools employed in the construction necessarily involve initial data modifications in the 'gluing zone', that is, they are quasi-local in the sense of footnote 13. On the other hand, in cases in which the construction of initial data 'from scratch' is actually needed (and this actually happens in concrete applications), the gluing engineering is not necessarily the appropriate approach. In any case, gluing techniques have proved to be very powerful and fruitful for demonstrating the existence of certain classes of spacetime solutions as well as of certain initial data sets, solving problems left open by conformal approaches. In this sense, it is fair to say that more investigations are needed to fully assess the status of gluing techniques in the context of the discussion about counterfactuals.[29]

So, from the philosophical point of view, evaluating counterfactuals—and in particular applying Maudlin's recipe—in general relativistic contexts is not impossible, but much more subtle than commonly thought. The upshot for Maudlin is that it might be much more difficult for him to show how (primitive) laws support counterfactuals in these GR cases.[30]

More broadly, there may well be GR situations where none of the (conformal decomposition, gluing, ...) techniques work—or even a more serious concern, where the very initial value formulation may simply not be appropriate. In particular, the Cauchy approach in GR provides an extraordinarily powerful scheme to understand the structure of the theory, for example to assess genericity issues in evolution problems (see footnote 28). However, it is not clear that such approach is always well suited for the dynamical study of an individual system in GR, namely through fixing its physical properties by a given prescription at some point in its evolution, due precisely to the difficulties to capture their physical content in the initial data (see for instance the example discussed in footnote 27 concerning the consistency check between observed gravitational wave signals and the reconstructed waveforms). This does not concern the 'actual' dynamics of the system itself, which is perfectly unambiguous when considering its whole evolution since its formation, but it rather concerns our 'description' of such dynamics when starting the modelling at an arbitrary given intermediate time. In this regard, whereas in the quantum context issues in the evaluation of counterfactuals can be of ontological nature,[31] in the GR setting these limitations rather have a more epistemic flavour.

Indeed, it is very important to stress that the non-local issues we have discussed in this article are circumscribed to our ability to construct initial data, more precisely to impose a prescribed physical content in the resolution process of the constraints.[32] They do not arise if one is somehow provided with appropriate initial data (already satisfying the constraints), since Bianchi identities

[29]In particular, gluing techniques directly impact Maudlin's step 2 and the ceteris paribus condition, through their capability to build new data from previous ones where non-local modifications, although not eliminated, are reduced from a fully global scale to a region around the original data (in other words, they are quasi-local). This is a remarkable result that illustrates the subtleties of ellipticity. Although the introduced modifications in the gluing zone do affect the evaluation of quasi-local quantities, therefore spoiling the full physical control of the initial data, it might well be the case that the consequent ambiguities prove to be harmless at an effective level. Understanding the underlying mechanism in such a picture, as well as its limits, constitutes a current challenge in the study of the strong field regime of spacetime dynamics in GR.

[30]This might slightly weaken his argument in favour of a primitivist conception about laws, since his strategy partly relies on showing the benefits of taking laws as primitive. However, as we have seen, his conception is not the only one facing these difficulties in the GR context.

[31]This of course depends on the interpretative stance on quantum theory.

[32]As we have argued, these issues are crucial to counterfactual evaluation, but also impact the analysis of specific GR problems.

then guarantee the exact satisfaction of the constraints along the whole evolution. In other words, in such case, one would never be confronted to the resolution of the constraints and their related non-local aspects highlighted here; the corresponding dynamical system itself—namely, spacetime—is 'unaware' of this problematics. It is in this sense that these specific non-local issues of GR present an epistemic nature.

Finally, let us note that difficulties for evaluating counterfactuals in the GR context may also have some implications for our understanding of the very explanatory power of GR (this is already suggested by the above mentioned potential difficulties for constructing appropriate initial data and therefore for extracting related predictions in GR). Indeed, explanatory power and counterfactuals are intuitively connected, as captured by Hall ([2015], p. 270), who argues for a Humean understanding of laws and counterfactuals precisely using the tools of the initial value formulation: 'the ability to provide sharp and determinate truth conditions for a wide range of counterfactuals is precisely what lends a good physical theory its explanatory power'. This understanding is perfectly valid for GR in the weak field regime, and accounts for much of its explanatory power in this domain. But in the strong field regime, this account of the explanatory power of GR may break down, because of the difficulties in evaluating counterfactuals we have discussed in this article. In spite of such difficulties, which are related to capturing sharp details in the initial data, the theory has proven remarkably predictive in the explored sector of the strong field regime, as illustrated by binary pulsar dynamics or gravitational wave emission. This hints (see also footnote 27) to the possibility of an underlying dynamical mechanism 'smoothing' or 'coarse graining' those sharp details in appropriate classes of initial data.[33] In this perspective, and to the extent that GR will remain valid in the entire strong field regime, the theory might actually require an alternative explanatory framework—perhaps one that does not rely on (or at least reformulate) certain aspects of the initial value formulation.

Acknowledgements

V.L. wishes to thank friends and colleagues at the University of Geneva and is grateful to the Swiss National Science Foundation for financial support (grant 169313). J.L.J. thoughtfully thanks M. Ansorg ('GR has teeth...') and S. Dain, together with E. Gourgoulhon and J.A. Valiente-Kroon, for generously sharing along the years their thorough insights into the initial value problem of general relativity. He also thanks the Arbaz community for their warm, always enlightening, hospitality.

José Luis Jaramillo
Institut de Mathématiques de Bourgogne (IMB)
UMR 5584, CNRS, Université de Bourgogne Franche-Comté
F-21000 Dijon, France
Jose-Luis.Jaramillo@u-bourgogne.fr

Vincent Lam
Chair of Excellence in Philosophy of Quantum Physics
Institut Néel & PPL, Université Grenoble Alpes

[33] GR already provides examples in this spirit in particular physical scenarios, for instance the 'effacement' property of the internal structure of compact self-gravitating bodies (Damour [1987]), the circularization of orbits by gravitational waves or the black hole 'no hair' results in late gravitational collapse. Note that such 'smoothing' is not in contradiction with the existence of initial data sets (of 'zero-measure' in the phase space of the theory) extremely sensitive to small modifications, as illustrated by 'critical surfaces' in the context of the discovery of critical phenomena in GR (see for example Gundlach and Martín-García [2007]).

F-38000 Grenoble, France
Vincent.Lam@univ-grenoble-alpes.fr
and
School of Historical and Philosophical Inquiry
The University of Queensland
St Lucia QLD 4072, Australia
v.lam@uq.edu.au

References

Bartnik, R. and Isenberg, J. [2004]: ‘The Constraint Equations’, in P. T. Chruściel and H. Friedrich (*eds*), *The Einstein Equations and the Large Scale Behavior of Gravitational Fields*, Basel: Birkhäuser, pp. 1-38.

Bigaj, T. [2006]: *Non-locality and Possible Worlds*, Heusenstamm: ontos.

Choquet-Bruhat, Y. and Geroch, R. [1969]: ‘Global aspects of the Cauchy problem in general relativity’, *Communications in Mathematical Physics*, 14, pp. 329-35.

Chruściel, P. T., Isenberg, J. and Pollack, D. [2005]: ‘Initial data engineering’, *Communications in Mathematical Physics*, 257, pp. 29-42.

Chruściel, P. T., Corvino, J. and Isenberg, J. [2011]: ‘Construction of n-body initial data sets in general relativity’, *Communications in Mathematical Physics*, 304, pp. 637-47.

Corvino, J. and Schoen, R. [2006]: ‘On the asymptotics for the vacuum Einstein constraint equations’, *Journal of Differential Geometry*, 73, pp. 185-217.

Curiel, E. [2015]: ‘If Metrical Structure Were Not Dynamical, Counterfactuals in General Relativity Would Be Easy’, <arxiv.org/pdf/1509.03866.pdf>.

Dain, S. [2006]: ‘Elliptic systems’, *Lecture Notes in Physics*, 692, pp. 117-39.

Damour, T. [1987]: ‘The problem of motion in Newtonian and Einsteinian gravity’, in S. Hawking and W. Israel (*eds*), *Three Hundred Years of Gravitation*, Cambridge: Cambridge University Press, pp. 128-98.

Friedman, J. L. [2004]: ‘The Cauchy Problem on Spacetimes That Are Not Globally Hyperbolic’, in P. T. Chruściel and H. Friedrich (*eds*), *The Einstein Equations and the Large Scale Behavior of Gravitational Fields*, Basel: Birkhäuser, pp. 331-46.

Geroch, R. [1969]: ‘Limits of spacetimes’, *Communications in Mathematical Physics*, 13, pp. 180-93.

Gundlach, C. and Martín-García, J. M. [2007]: ‘Critical phenomena in gravitational collapse’, *Living Reviews In Relativity*, 10, <www.livingreviews.org/lrr-2007-5>.

Hall, N. [2015]: ‘Humean reductionism about laws of nature’, in B. Loewer and J. Schaffer (*eds*), *A Companion to David Lewis*, Chichester: Wiley, pp. 262-77.

Healey, R. [2009]: ‘Causation in Quantum Mechanics’, in H. Beebee, C. Hitchcock and P. Menzies *(eds)*, *The Oxford Handbook of Causation*, Oxford: Oxford University Press, pp. 673-86.

Isenberg, J. [2014]: ‘The Initial Value Problem in General Relativity’, in A. Ashtekar and V. Petkov (*eds*), *Springer Handbook of Spacetime*, Dordrecht: Springer, pp. 303-21.

Lewis, D. [1973]: *Counterfactuals*, Oxford: Blackwell.

Lewis, D. [1979]: ‘Counterfactual Dependence and Time’s Arrow’, *Noûs*, 13, pp. 455-76.

Maudlin, T. [2007]: *The Metaphysics within Physics*, Oxford: Oxford University Press.

Norton, J. [1992]: ‘Einstein, Nordström and the early Demise of Lorentz-covariant Scalar Theories of Gravitation’, *Archive for History of Exact Sciences*, 45, pp.17-94.

Pfeiffer, H. P. and York Jr., J. W. [2003]: ‘Extrinsic curvature and the Einstein constraints’, *Physical Review D*, 67, 044022.

Rácz, I. [2016]: ‘Constraints as evolutionary systems’, *Classical and Quantum Gravity*, 33, 015014.

Rueede, C. and Straumann, N. [1997]: ‘On Newton-Cartan cosmology’, *Helvetica Physica Acta*, 70, pp. 318-35.

Szabados, L. [2009]: ‘Quasi-Local Energy-Momentum and Angular Momentum in General Relativity’, *Living Reviews In Relativity*, 12, <www.livingreviews.org/lrr-2009-4>.

York Jr., J. W. [1999]: ‘Conformal “thin-sandwich” data for the initial-value problem of general relativity’, *Physical Review Letters*, 82, pp. 1350-3.