The relativity and equivalence principles for self-gravitating systems

David Wallace*

March 25, 2009

Abstract

I criticise the view that the relativity and equivalence principles are consequences of the small-scale structure of the metric in general relativity, by arguing that these principles also apply to systems with non-trivial self-gravitation and hence non-trivial spacetime curvature (such as black holes). I provide an alternative account, incorporating aspects of the criticised view, which allows both principles to apply to systems with self-gravity.

1 Introduction: two principles

The relativity principle — the observation that the laws of physics are the same in two reference frames in constant motion with respect to one another — is by now very well understood at least from a mathematical point of view. It can be understood as entailed by certain symmetries of the laws of physics: the Gallilei symmetry group in pre-relativistic physics, the Poincaré group in special relativity. Alternatively, and perhaps equivalently, it can be understood as a consequence of the structure of spacetime: in either neo-Newtonian[1] or Minkowski spacetime, boosts between reference frames are automorphisms of the background spacetime structure and so have no physically detectable consequences.

This raises a question: our world is apparently correctly described (or at any rate *better* described) by general relativity rather than by special relativity (let alone Newtonian physics). So the laws of physics do not have the Poincarè group as a symmetry (indeed, it is not even clear what that would mean), and the spacetime structure is described by a metric which does not, in general, have any automorphisms. So why is the relativity principle empirically correct, if its theoretical underpinnings are not?

The conventional wisdom appears to be that the relativity principle is recovered as a result holding in regions which small enough that spacetime curvature

---

*Balliol College, Oxford University; email david.wallace@philosophy.ox.ac.uk
[1]See (e.g.) chapter 2 of Earman (1989) for an account of neo-Newtonian spacetime.

can be neglected. In terms of the metric: in a sufficiently small spacetime patch, curvature is negligible so the metric may be approximated as Minkowskian. In terms of the laws of physics: in a sufficiently small spacetime patch, the laws of special relativity hold to a high level of accuracy. (See the work of Harvey Brown, and in particular Brown (2005, pp. 169–172) for a particularly clear statement of this position.)

Just as the relativity principle was historically crucial in the development of special relativity, so the equivalence principle played a crucial part in developing general relativity. A natural way to state it is as the claim that a system falling freely in a uniform gravitational field will behave in exactly the same way as an isolated system. In Newtonian physics, this makes literal sense, and can be understood as the consequence of a symmetry of the theory: if we take the theory to be specified by a collection of point particles and by a gravitational potential $V(\mathbf{x}, t)$, where $V(\mathbf{x}, t)$ satisfies the Poisson equation and the particles obey Newton's laws (possibly with other distance-dependent or contact forces present), then the process of giving the masses uniform acceleration $\mathbf{a}(t)$ and adding a term $\mathbf{a}(t) \cdot \mathbf{x}$ to the potential is a process which takes solutions to solutions. While there is no geometric equivalent of this dynamical symmetry as long as we formulate Newtonian physics on neo-Newtonian spacetime, the very presence of this mismatch between dynamical and spacetime symmetries points to the existence of a more natural geometric arena for Newtonian gravity: Newton-Cartan spacetime, the automorphisms of which include arbitrary spatially constant accelerations.

It is somewhat harder to make sense of the equivalence principle in this way in general relativity, where there is no direct analogue to the gravitational potential. But a natural point of connection is the idea that physics should be the same in any freely falling reference frame. Once that is understood, we are faced with a similar problem as for the relativity principle — how to define "reference frame" in a setting which has spacetime curvature — and a similar answer is available. Namely, the equivalence principle (says the conventional wisdom) is a corollary of the fact that sufficiently small regions of spacetime may be treated as flat (and hence isomorphic to regions of Minkowski spacetime). In particular, since in a sufficiently thin tube around an arbitrary geodesic the spacetime metric may be approximated as flat, physics within that tube ought to be indistinguishable from physics within a similarly shaped tube in Minkowski spacetime (or indeed in intergalactic space).

So we have a rather straightforward story:the relativity and equivalence principles both hold true, in general relativity, because of the small-scale behaviour of the metric, and in particular because in any metric, a sufficiently small spacetime patch (or a sufficiently thin spacetime tube around some geodesic) may be idealised as metrically flat. Neither principle, therefore, is concerned with the *large*-scale behaviour of the metric in solutions to the Einstein field equations; neither can apply in situations where curvature cannot be neglected.

The point of this paper is to challenge this story: it is not right, or at any rate it is not at all the whole truth, and there is an alternative and preferable account. In section 2 I will give examples to show just why there must be more

to say; in sections 3–5 I will try to say it.

2   Galileo's Black Hole

The folklore of physics (correctly, so far as I know) attributes the first statement of the relativity principle to Galileo, in his famous thought-experiment:

> Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you there some flies, butterflies, and other small flying animals. Have a large bowl of water with some fish in it; hang up a bottle that empties drop by drop into a wide vessel beneath it.

> With the ship standing still, observe carefully how the little animals fly with equal speed to all sides of the cabin. The fish swim indifferently in all directions; the drops fall into the vessel beneath; and, in throwing something to your friend, you need to throw it no more strongly in one direction than another, the distances being equal; jumping with your feet together, you pass equal spaces in every direction.

> When you have observed all of these things carefully (though there is no doubt that when the ship is standing still eveything must happen this way), have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still. (Galileo (1960))

Tragically (if unsurprisingly), Galileo did not live to see the development of general relativity. If he had, who can doubt that he would have quickly penned a sequel, which no doubt would have become known as *Galileo's black hole*:

> Put yourself together with some friend in orbit around some large black hole, and have with you there some planetoids, interstellar dust, and other solid matter. Have a long rope with a clock at one end of it; place a light source into an orbit below yours. With the black hole standing still, observe carefully how the redshift from the light source is of equal magnitude at all points in its orbit. The clock slows equally whatever side of the hole it is lowered towards; and, in allowing the solid matter to fall inward and form an accretion disk, the disk forms no more strongly on one side than another; observing the high-energy jets above and below its plane, they are as like to emerge in one direction as in the other.

> When you have observed all these things carefully (though there is no doubt that when the black hole is standing still everything must happen in this way), have the hole proceed with any speed you like,

so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the hole was moving or standing still.

The relativity principle, in other words, had better apply to all manner of systems, including those which are strongly self-gravitating. More to the point (for what does Nature care what "had better" be true?), there is abundant evidence that it *does* apply to such systems. Observations of black holes and neutron stars are commonplace in high-energy astrophysics these days, and astrophysicists do not even consider the "absolute velocity" of the system they study (whatever that would be) in their analysis. And at a more mundane level, even planets have non-trivial self-gravity, and the relativity principle manifestly applies to that self-gravity itself. Objects fall under Earth's gravity in just the same way in summer as in winter, despite the ~ 60 kilometre-per-second velocity difference.

But the analysis I gave of the relativity principle in section 1 cannot possibly apply to self-gravitating systems. For that analysis required the principle to apply only in regions where the metric could be idealised as flat, and what is "a region where the metric can be idealised as flat", if not a region in which gravitational phenomena can be idealised away? (Or, put more starkly, regions of spacetime in which the metric is approximately flat do not as a rule contain black holes.)

And the same is true of the equivalence principle. The Earth-Moon system is in orbit around the Sun, and thus moving freely in the Sun's gravitation field, yet the Moon orbits the Earth, and objects fall under Earth's gravity, just as they would were the Sun not there.[2] Drop a star — or even a neutron star or black hole — deep into the gravitational field of a supermassive black hole. At least until tidal forces become significant, the star's own physics — including the star's gravitational physics — will show no sign that it is not in open space far from the supermassive hole.

So: physically speaking both the equivalence principle and the relativity principle make sense not only in flat regions of spacetime but for systems with significant self-gravity. In the next section I will consider how to make sense of this fact mathematically speaking, and thus show what makes it correct.

3 Poincaré transformations of gravitating systems

Consider the spacetime of a single stationary black hole. Far from the event horizon of that black hole, spacetime is flat to an arbitrarily good approximation; in such distant regions, it makes as much sense to talk about "inertial observers"

---

[2] To a good approximation, at any rate; the only observable effect of which I am aware (other than astronomical observations) is the monthly variation in the strength of the tides as Sun and Moon move in and out of alignment.

as it does in Minkowski spacetime. (At least, it does so provided we restrict our attention to timescales on which the "arbitrarily good approximation" we have made really is good enough: on sufficiently long timescales, all observers eventually move on trajectories influenced by the black hole's presence.)

There is a preferred set of such observers: those whose distance from the black hole remains constant. Their worldlines (again, on sufficiently short timescales) coincide with the integral curves of the black hole's timelike Killing field; it makes sense to regard them as stationary relative to the black hole. As such, observers moving at velocity $\mathbf{v}$ with respect to the stationary observers can reasonably be thought of as moving at velocity $\mathbf{v}$ with respect to the black hole.

Now consider a diffeomorphism of the spacetime which takes the stationary observers onto the velocity-$\mathbf{v}$ observers. By analogy with special relativity, it is tempting to regard this diffeomorphism as representing a boost of the black hole by velocity $\mathbf{v}$ (and hence to regard a boost as defined only by its long-distance behaviour).

Should we yield to this temptation? As is now well recognised, we can only really get at the empirical consequences of symmetry transformations when we consider transformations of a subsystem of the universe — such as setting Gallileo's ship into motion — and not the whole universe. (See Greaves and Wallace (2009) for a more detailed discussion.) So to understand physically what makes the diffeomorphism we have constructed "count" as a boost transformation, we need to add something into our model other than the black hole.

This can be done straightforwardly, at least approximately speaking. Suppose that Σ_1 and Σ_2 are non-intersecting spacelike hypersurfaces[3], and let us restrict our attention to the region of spacetime between the two. Sufficiently far from the black hole, there will be a tube of spacetime connecting the two surfaces the metric of which is Minkowskian to as good an approximation as we like; as such, we should be able to excise the interior of that region and slot in *another* black hole, or more accurately a tubular region cut from another spacetime containing one black hole. If the tube's radius is large compared to the black hole mass[4] then the discontinuity at the boundary will be small, and it is reasonable to suppose that the resulting spacetime is close to an exact solution of the field equations; certainly this is the routine assumption tacit in applications of idealised black-hole solutions to actual black holes in our universe.

To define what it means to perform a diffeomorphism on only part of this system — that is, to perform it one one of the black holes and not on the other — is a bit more complicated. Intuitively what we want to do is to allow the diffeomorphism to act only on the points within the surface defining a particular isolated subsystem (one of the black holes, in this case). But defining things in that way will lead to problems at the boundary of the subsystem: on some regions around the boundary the metric and matter fields will be multiply defined, on others they will not be defined at all. (For instance, suppose we just

---

[3]More accurately: they need to be spacelike *far from the event horizon* (this kind of awkward detail is the price I pay for using black holes in my example!)

[4]Large in natural units, that is: $GM/c^2 \ll R$.

consider a one-dimensional manifold diffeomorphic to the real line, coordinatise it by real numbers, and try to apply a diffeomorphism only to the subset $(0, 1)$. If the diffeomorphism (in these coordinates) is given by $f(x) = x + \epsilon$ (with $\epsilon < 1$) then applying it only to $(0, 1)$ will leave the metric and matter fields multiply defined in $(1, 1 + \epsilon)$ and undefined in $(0, \epsilon)$.

To give a definition that gets around this problem, I need a bit of formalism. Firstly, I define a dynamically possible spacetime (or just 'spacetime', for short) as a triple $\langle \mathcal{M}, g, \psi \rangle$, where $\mathcal{M}$ is a four-manifold, and g and ψ are metric and matter fields respectively on $\mathcal{M}$ jointly satisfying the dynamical equations of the system. (ψ should be understood as a placeholder for whatever matter fields are actually of interest). I will say that some subset of the manifold is a *temporal segment* of the spacetime iff it consists of all the points in the manifold between some non-intersecting hypersurfaces Σ_1 and Σ_2. (The intuitive idea here would require the hypersurfaces to be spacelike at least in "well-behaved" regions of spacetime, but I refrain from attempting to formalise this.)

Now suppose that S is a closed surface in some temporal segment τ of a spacetime $\langle \mathcal{M}, g, \psi \rangle$. I will say that the interior of S is an isolated subsystem of the temporal segment if there exists a temporal segment τ' of some spacetime $\langle \mathcal{M}', g, \psi' \rangle$ and a closed surface S' in that spacetime, such that

(i) The interior of S' is approximately diffeomorphic to the interior of S (that is, there is a diffeomorphism between the two which takes the metric and matter fields inside S onto the metric and matter fields inside S';

(ii) That part of τ' not in the interior of S' is approximately flat and approximately empty of matter.

It follows, of course, that the spacetime region in the vicinity of S must also be approximately flat and empty. And it should, I hope, be clear that there are surfaces surrounding the two black holes in the previous example are each isolated subsystems.

A diffeomorphism of a subsystem can now be defined as follows. Suppose $\mathcal{M}, \tau, g, \psi, S$ and $\mathcal{M}', \tau', g', \psi', S'$ are as stated above and that there is a diffeomorphism φ from the interior of S onto the interior of S' which maps $g|_S$ to $g'|_{S'}$ and $\psi|_S$ to $\psi'|_{S'}$. Now let χ be some diffeomorphism of $\mathcal{M}'$ such that τ' remains approximately flat and empty of matter after the diffeomorphism is performed. (Note that we transform matter and metric fields by χ, but leave S' and the surfaces defining τ' alone). Then we can define new matter and metric fields $\tilde{g}$, $\tilde{\psi}$ on τ as follows: they coincide with the old matter and metric fields outside S, and inside S they are equal to the pull-back under φ of the χ-transformed matter and metric fields within S'. That is:

$$\tilde{g} = \varphi^* \chi_* g' \text{ (inside } S)$$

$$\tilde{g} = g \text{ (outside } S) \tag{1}$$

and similarly for $\tilde{\psi}$.

Our particular interest is in those χ which are Poincaré transformations (that is, automorphisms of the flat metric) outside S'. Precisely because these transformations leave the flat metric invariant, they induce transformations of τ which preserve the boundary conditions of S, and hence which take solutions of the field equations to solutions of the field equations. (Since clearly the field equations are satisfied by $\tilde{g}, \tilde{\psi}$ both inside and outside S, preserving the boundary conditions at S is a necessary and sufficient condition for the induced transformation to be a solution of the dynamical equations.)

Furthermore, if we consider the worldlines of inertially moving observers just inside and just outside S, it is clear that the effect of the transformation is to induce a Poincaré transformation of those inside S relative to those outside S. This suffices to give physical content to my earlier claim that the general-relativistic expression of a Lorentz boost of a black hole is a diffeomorphism which reduces to a Lorentz boost in the flat-spacetime region far from the hole. More generally, indeed, transformations of isolated systems which are asymptotically Poincaré do indeed represent Poincaré shifts of that system relative to any other systems which are present.

(In fact, this analysis is a special case of a more general principle that the physically significant symmetries of a system are those which (a) do not vanish at the system's boundary, but (b) leave the boundary conditions of the system invariant. See Greaves and Wallace (2009) for a more systematic working-out of this idea.)

4 The relativity and equivalence principles

We are now in a position to answer one of this paper's main questions: namely, how to understand the relativity principle in general relativistic systems. And despite the technical details required to define subsystem transformations in general relativity, the overall story is fairly simple. Isolated systems, as a matter of dynamics, lead to metrics which are asymptotically flat, and therefore can be approximated as flat at large but finite distances from the centre of the system. Hence in a spacetime which consists of a discrete number of isolated systems embedded in an otherwise (approximately) flat spacetime, applying a diffeomorphism to an isolated subsystem which is Poincaré (that is, is an automorphism of the flat matric) on the subsystem's boundary will generate a new spacetime in which

(a) the matter and metric fields still (approximately) satisfy the dynamical equations;

(b) the subsystem has been translated and/or boosted and/or rotated relative to the other subsystems (in a way which is perfectly detectable if they are allowed to interact at some later time);

(c) physical processes in the subsystem are unaffected.

The problem with this account, as it stands, is that it assumes that spacetime is approximately flat in the large, which it certainly is not.[5] On the other hand, it is often possible in practice to find regions of spacetime both large enough to contain several isolated systems, and small enough to be treated as approximately flat.

Consider, for instance, physics in the vicinity of some supermassive black hole of, say, a billion solar masses. Because the tidal forces generated by such a large black hole are quite small (provided we are sufficiently far from the singularity), there will be quite large regions of spacetime in the vicinity of the hole such that we can treat them as approximately flat.

In particular, we should be able to find some length L such that (a) the spacetime of the supermassive black hole can be treated as approximately flat in regions of size L, but (b) L is much larger than the effective size of (say) an ordinary stellar-mass black hole. If this is the case, we will be able to embed several such stellar-mass black holes (or other isolated systems of similar or lesser mass) into such a region, and apply the arguments of the previous section to make sense of Poincaré transformations of one such isolated system with respect to another. (And nothing in this analysis relies on the choice of a supermassive black hole as the source of the low-curvature background metric: the metric generated by the galaxy, or indeed the cosmological metric, would do as well.)

And in fact, this same analysis accounts for why the equivalence principle holds for gravitating systems. Any isolated gravitational system which is approximately flat outside a sphere of radius[6] $\sim l$ can be embedded into any background metric where the curvature (in the relevant region) can be neglected on scales of $\sim L$, where $L \gg l$: just select a sphere of radius l in the background metric and in the metric of the isolated system, and excise the interior of the former and the exterior of the latter. Because the metric in both cases is approximately flat at the boundary, the two surfaces will be diffeomorphic and so we can find a scheme to paste the two together whilst preserving the boundary conditions.

For instance, consider the geodesic of a test particle falling into a supermassive black hole. For given time τ and sufficiently small length L, there will be a tube of length τ and radius L around the geodesic in which deviations from flatness are trivial. Since the geodesic is straight, the region will be (approximately) diffeomorphic to a rectilinear tube in Minkowski spacetime. Given a black hole, or other isolated system, whose metric is approximately flat on scales significantly smaller than L, we can paste that system into the tube around the geodesic and so obtain a solution to the dynamical equations. It follows that the behaviour of the subsystem is unaffected by its passage towards the black hole (until tidal forces become so high as to be relevant on the lengthscales of the subsystem). That is, it follows that the equivalence principle applies to the

---

[5]Current cosmological observations are consistent with the possibility that *space* is flat in the large, but this is a much weaker claim.

[6]Since, for self-gravitating systems, measuring the distance from boundary to centre will mean penetrating the region of high curvature and can (e. g. , in the case of black holes) give misleading results, strictly I ought to say "approximately flat outside a sphere of area $\sim l^2$".

subsystem.

5 Conclusions

We can coherently talk about isolated systems in general relativity because, as a matter of dynamics, there exist a large number of solutions to the equations — including ones which represent stars, planets, black holes, etc, as well as interacting sets of these — where the curvature and matter are concentrated in some finite region and far outside that region the spacetime is approximately empty and flat. This allows us to paste such solutions together, to form regions of spacetime consisting of a number of isolated subsystems embedded in approximately flat spacetime. Because of the Poincaré symmetry of flat spacetime, we can perform a Poincaré transformation on one of the subsystems without violating the boundary conditions between subsystems; hence, the relativity principle applies for collections of such subsystems.

In turn, regions of effectively flat spacetime can always be found in a given spacetime, provided we are prepared to make those regions sufficiently small. If "sufficiently small" is nonetheless large compared to the effective size of the subsystems we are interested in, then (a) we can apply the above argument for the relativity principle to isolated systems in a curved spacetime; (b) we can embed such systems in any such effectively flat region without affecting their internal dynamics, since their Minkowski boundary conditions are compatible with any region flat on sufficiently large lengthscales.

In effect, then, the equivalence principle applies in general relativity because the metric of isolated systems at sufficiently *large* distances is the same as the metric of any system at sufficiently *small* distances. The relativity principle applies because, in addition to this, that metric has the Poincaré group as a symmetry group.

I began this paper by criticizing the 'conventional wisdom' that the relativity and equivalence principles apply to general relativity because of the flatness of the metric at sufficiently short lengthscales. I hope that I have ended by showing that this is at most half the story: the flatness of the metric of certain systems at sufficiently large lengthscales has an equally important role. There is therefore an important dynamical aspect to both principles which does not seem to have been widely recognised: whether or not the small-scale behaviour of the metric should be understood as 'mere kinematics', it is a nontrivial dynamical fact that those solutions of the Einstein field equations appropriate for the description of isolated systems are asymptotically flat.

Acknowledgements

References

Brown, H. R. (2005). *Physical Relativity*. Oxford: Oxford University Press.

Earman, J. (1989). *World Enough and Space-Time*. Cambridge, Massachusetts: MIT Press.

Galileo (1960). *Dialogue Concerning Two Chief World Systems*. Berkeley, California: University of California Press. Translated by S. Drake.

Greaves, H. and D. Wallace (2009). Empirical consequences of symmetries: a new framework. Forthcoming.