Massachusetts Institute of Technology
Department of Physics

Physics 8.962 Spring 2002

Symmetry Transformations, the Einstein-Hilbert Action, and Gauge Invariance

1 Introduction

Action principles are widely used to express the laws of physics, including those of general relativity. For example, freely falling particles move along geodesics, or curves of extremal path length.

Symmetry transformations are changes in the coordinates or variables that leave the action invariant. It is well known that continuous symmetries generate conservation laws (Noether's Theorem). Conservation laws are of fundamental importance in physics and so it is valuable to investigate symmetries of the action.

It is useful to distinguish between two types of symmetries: **dynamical** symmetries corresponding to some inherent property of the matter or spacetime evolution (e.g. the metric components being independent of a coordinate, leading to a conserved momentum one-form component) and **nondynamical** symmetries arising because of the way in which we formulate the action. Dynamical symmetries constrain the solutions of the equations of motion while nondynamical symmetries give rise to mathematical identities. These notes will consider both.

An example of a nondynamical symmetry is the parameterization-invariance of the path length, the action for a free particle:

$$S[x^\mu(\tau)] = \int_{\tau_1}^{\tau_2} L_1\left(x^\mu(\tau), \dot{x}^\mu(\tau), \tau\right)\, d\tau = \int_{\tau_1}^{\tau_2} \left[g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\right]^{1/2} d\tau\ . \qquad (1)$$

This action is invariant under arbitrary reparameterization $\tau \to \tau'(\tau)$, implying that any solution $x^\mu(\tau)$ of the variational problem $\delta S = 0$ immediately gives rise to other solutions

$y^\mu(\tau) = x^\mu(\tau'(\tau))$. Moreover, even if the action is not extremal with Lagrangian L_1 for some (non-geodesic) curve $x^\mu(\tau)$, it is still invariant under reparameterization of that curve.

There is another nondynamical symmetry of great importance in general relativity, coordinate-invariance. Being based on tensors, equations of motion in general relativity hold regardless of the coordinate system. However, when we write an *action* involving tensors, we must write the *components* of the tensors in some basis. This is because the calculus of variations works with functions, e.g. the components of tensors, treated as spacetime fields. Although the values of the fields are dependent on the coordinate system chosen, the action must be a scalar, and therefore invariant under coordinate transformations. This is true whether or not the action is extremized and therefore it is a nondynamical symmetry.

Nondynamical symmetries give rise to special laws called identities. They are distinct from conservation laws because they hold whether or not one has extremized the action.

The material in these notes is generally not presented in this form in the GR textbooks, although much of it can be found in Misner et al if you search well. Although these symmetry principles and methods are not needed for integrating the geodesic equation, they are invaluable in understanding the origin of the contracted Bianchi identities and stress-energy conservation in the action formulation of general relativity. More broadly, they are the cornerstone of gauge theories of physical fields including gravity.

Starting with the simple system of a single particle, we will advance to the Lagrangian formulation of general relativity as a classical field theory. We will discover that, in the field theory formulation, the contracted Bianchi identities arise from a non-dynamical symmetry while stress-energy conservation arises from a dynamical symmetry. Along the way, we will explore Killing vectors, diffeomorphisms and Lie derivatives, the stress-energy tensor, electromagnetism and charge conservation. We will discuss the role of continuous symmetries (gauge invariance and diffeomorphism invariance or general covariance) for a simple model of a relativistic fluid interacting with electromagnetism and gravity. Although this material goes beyond what is presented in lecture, it is not very advanced mathematically and it is recommended reading for students wishing to understand gauge symmetry and the parallels between gravity, electromagnetism, and other gauge theories.

2 Parameterization-Invariance of Geodesics

The parameterization-invariance of equation (1) may be considered in the broader context of Lagrangian systems. Consider a system with n degrees of freedom — the generalized coordinates q^i — with a parameter t giving the evolution of the trajectory in configuration space. (In eq. 1, q^i is denoted x^μ and t is τ.) We will drop the superscript on q^i when it is clear from the context.

Theorem: If the action $S[q(t)]$ is invariant under the infinitesimal transformation $t \to t + \epsilon(t)$ with $\epsilon = 0$ at the endpoints, then the Hamiltonian vanishes identically.

The proof is straightforward. Given a parameterized trajectory $q^i(t)$, we define a new parameterized trajectory $\bar{q}(t) = q(t + \epsilon)$. The action is

$$S[q(t)] = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt \ . \tag{2}$$

Linearizing $\bar{q}(t)$ for small ϵ,

$$\bar{q}(t) = q + \dot{q}\epsilon \ , \quad \frac{d\bar{q}}{dt} = \dot{q} + \frac{d}{dt}(\dot{q}\epsilon) \ .$$

The change in the action under the transformation $t \to t + \epsilon$ is, to first order in ϵ,

$$\begin{aligned} S[q(t+\epsilon)] - S[q(t)] &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial t}\epsilon + \frac{\partial L}{\partial q^i}\dot{q}^i\epsilon + \frac{\partial L}{\partial \dot{q}^i}\frac{d}{dt}(\dot{q}^i\epsilon) \right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{dL}{dt}\epsilon + \left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i \right) \frac{d\epsilon}{dt} \right] dt \\ &= [L\epsilon]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L \right) \frac{d\epsilon}{dt}\, dt \ . \end{aligned} \tag{3}$$

The boundary term vanishes because $\epsilon = 0$ at the endpoints. Parameterization-invariance means that the integral term must vanish for arbitrary $d\epsilon/dt$, implying

$$H \equiv \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L = 0 \ . \tag{4}$$

Nowhere did this derivation assume that the action is extremal or that $q^i(t)$ satisfy the Euler-Lagrange equations. Consequently, equation (4) is a nondynamical symmetry.

The reader may easily check that the Hamiltonian H_1 constructed from equation (1) vanishes identically. This symmetry does not mean that there is no Hamiltonian formulation for geodesic motion, only that the Lagrangian L_1 has non-dynamical degrees of freedom that must be eliminated before a Hamiltonian can be constructed. (A similar circumstance arises in non-Abelian quantum field theories, where the non-dynamical degrees of freedom are called Faddeev-Popov ghosts.) This can be done by replacing the parameter with one of the coordinates, reducing the number of degrees of freedom in the action by one. It can also be done by changing the Lagrangian to one that is no longer invariant under reparameterizations, e.g. $L_2 = \frac{1}{2}g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$. In this case, $\partial L_2/\partial\tau = 0$ leads to a *dynamical* symmetry, $H_2 = \frac{1}{2}g^{\mu\nu}p_\mu p_\nu =$ constant along trajectories which satisfy the equations of motion.

The identity $H_1 = 0$ is very different from the conservation law $H_2 =$ constant arising from a time-independent Lagrangian. The conservation law holds only for solutions of the equations of motion; by contrast, when the action is parameterization-invariant, $H_1 = 0$ holds for any trajectory. The nondynamical symmetry therefore does not constrain the motion.

3 Generalized Translational Symmetry

Continuing with the mechanical analogy of Lagrangian systems exemplified by equation (2), in this section we consider translations of the configuration space variables. If the Lagrangian is invariant under the translation $q^i(t) \to q^i(t) + a^i$ for constant a^i, then $p_i a^i$ is conserved along trajectories satisfying the Euler-Lagrange equations. This well-known example of translational invariance is the prototypical dynamical symmetry, and it follows directly from the Euler-Lagrange equations. In this section we generalize the concept of translational invariance by considering spatially-varying shifts and coordinate transformations that leave the action invariant. Along the way we will introduce several important new mathematical concepts.

In flat spacetime it is common to perform calculations in one reference frame with a fixed set of coordinates. In general relativity there are no preferred frames or coordinates, which can lead to confusion unless one is careful. The coordinates of a trajectory may change either because the trajectory has been shifted or because the underlying coordinate system has changed. The consequences of these alternatives are very different: under a coordinate transformation the Lagrangian is a scalar whose form and value are unchanged, while the Lagrangian can change when a trajectory is shifted. The Lagrangian is always taken to be a scalar in order to ensure local Lorentz invariance (no preferred frame of reference). In this section we will carefully sort out the effects of both shifting the trajectory and transforming the coordinates in order to identify the underlying symmetries. As we will see, conservation laws arise when shifting the trajectory is equivalent to a coordinate transformation.

We consider a general, relativistically covariant Lagrangian for a particle, which depends on the velocity, the metric, and possibly on additional fields:

$$S[x(\tau)] = \int_{\tau_1}^{\tau_2} L(g_{\mu\nu}, A_\mu, \ldots, \dot{x}^\mu)\, d\tau \ . \tag{5}$$

Note that the coordinate-dependence occurs in the fields $g_{\mu\nu}(x)$ and $A_\mu(x)$. An example of such a Lagrangian is

$$L = \frac{1}{2} g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu + qA_\mu \dot{x}^\mu \ . \tag{6}$$

The first piece is the quadratic Lagrangian L_2 that gives rise to the geodesic equation. The additional term gives rise to a non-gravitational force. The Euler-Lagrange equation for this Lagrangian is

$$\frac{D^2x^\mu}{d\tau^2} = qF^\mu{}_\nu \frac{dx^\nu}{d\tau} \ , \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu \ . \tag{7}$$

We see that the non-gravitational force is the Lorentz force for a charge q, assuming that the units of the affine parameter τ are chosen so that $dx^\mu/d\tau$ is the 4-momentum (i.e. $md\tau$ is proper time for a particle of mass m). The one-form field $A_\mu(x)$ is the

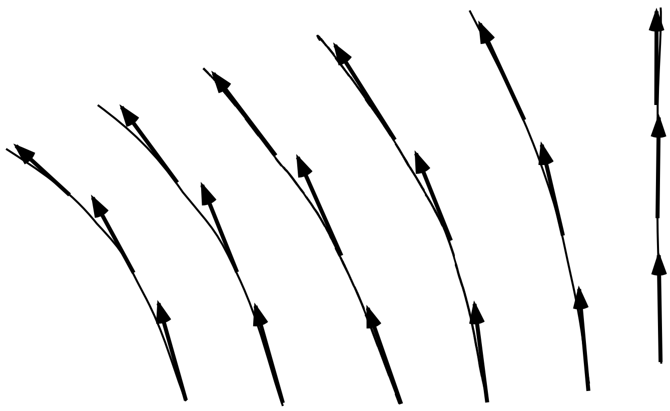

Figure 1: A vector field and its integral curves.

electromagnetic potential. We will retain the electromagnetic interaction term in the Lagrangian in the presentation that follows in order to illustrate more broadly the effects of symmetry.

Symmetry appears only when a system is changed. Because L is a scalar, coordinate transformations for a fixed trajectory change nothing and therefore reveal no symmetry. So let us try changing the trajectory itself. Keeping the coordinates (and therefore the metric and all other fields) fixed, we will shift the trajectory along the integral curves of some vector field $\xi^\mu(x)$. (Here $\vec{\xi}$ is any vector field.) As we will see, a vector field provides a one-to-one mapping of the manifold back to itself, providing a natural translation operator in curved spacetime.

Figure 1 shows a vector field and its integral curves $x^\mu(\lambda, \tau)$ where τ labels the curve and λ is a parameter along each curve. Any vector field $\vec{\xi}(x)$ has a unique set of integral curves whose tangent vector is $\partial x^\mu/\partial\lambda = \xi^\mu(x)$. If we think of $\vec{\xi}(x)$ as a fluid velocity field, then the integral curves are streamlines, i.e. the trajectories of fluid particles.

The integral curves of a vector field provide a continuous one-to-one mapping of the manifold back to itself, called a pushforward. (The mapping is one-to-one because the integral curves cannot intersect since the tangent is unique at each point.) Figure 2 illustrates the pushforward. This mapping associates each point on the curve $x^\mu(\tau)$ with a corresponding point on the curve $y^\mu(\tau)$. For example, the point P_0 ($\lambda = 0, \tau = 3$) is mapped to another point P ($\lambda = 1, \tau = 3$). The mapping $x \to y$ is obtained by

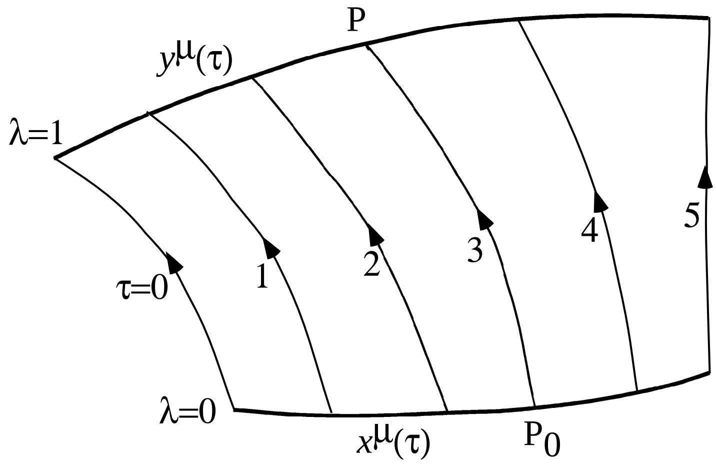

Figure 2: Using the integral curves of a vector field to shift a curve $x^\mu(\tau)$ to a new curve $y^\mu(\tau)$. The shift, known as a pushforward, defines a continuous one-to-one mapping of the space back to itself.

integrating along the vector field $\vec{\xi}(x)$:

$$\frac{\partial x^\mu}{\partial \lambda} = \xi^\mu(x)\ ,\quad x^\mu(\lambda = 0, \tau) \equiv x^\mu(\tau)\ ,\quad y^\mu(\tau) \equiv x^\mu(\lambda = 1, \tau)\ . \tag{8}$$

The shift amount $\lambda = 1$ is arbitrary; any shift along the integral curves constitutes a pushforward. The inverse mapping from $y \to x$ is called a pullback.

The pushforward generalizes the simple translations of flat spacetime. A finite translation is built up by a succession of infinitesimal shifts $y^\mu = x^\mu + \xi^\mu d\lambda$. Because the vector field $\vec{\xi}(x)$ is a tangent vector field, the shifted curves are guaranteed to reside in the manifold.

Applying an infinitesimal pushforward yields the action

$$S[x(\tau) + \xi(x(\tau))d\lambda] = \int_{\tau_1}^{\tau_2} L(g_{\mu\nu}(x + \xi d\lambda), A_\mu(x + \xi d\lambda), \dot{x}^\mu + \dot{\xi}^\mu d\lambda)\, d\tau\ . \tag{9}$$

This is similar to the usual variation $x^\mu \to x^\mu + \delta x^\mu$ used in deriving the Euler-Lagrange equations, except that ξ is a field defined everywhere in space (not just on the trajectory) and we do not require $\xi = 0$ at the endpoints. Our goal here is not to find a trajectory that makes the action stationary; rather it is to identify symmetries of the action that result in conservation laws.

We will ask whether applying a pushforward to one solution of the Euler-Lagrange equations leaves the action invariant. If so, there is a dynamical symmetry and we will obtain a conservation law. Note that our shifts are more general than the uniform translations and rotations considered in nonrelativistic mechanics and special relativity (here the shifts can vary arbitrarily from point to point, so long as the transformation has an inverse), so we expect to find more general conservation laws.

On the face of it, any pushforward changes the action:

$$S[x(\tau)+\xi(x(\tau))d\lambda] = S[x(\tau)]+d\lambda \int_{\tau_1}^{\tau_2} \left[\frac{\partial L}{\partial g_{\mu\nu}}(\partial_\alpha g_{\mu\nu})\xi^\alpha + \frac{\partial L}{\partial A_\mu}(\partial_\alpha A_\mu)\xi^\alpha + \frac{\partial L}{\partial \dot{x}^\mu}\frac{d\xi^\mu}{d\tau}\right] d\tau \ . \tag{10}$$

It is far from obvious that the term in brackets ever would vanish. However, we have one more tool to use at our disposal: coordinate transformations. Because the Lagrangian is a scalar, we are free to transform coordinates. In some circumstances the effect of the pushforward may be eliminated by an appropriate coordinate transformation, revealing a symmetry.

We consider transformations of the coordinates $x^\mu \to \bar{x}^\mu(x)$, where we assume this mapping is smooth and one-to-one so that $\partial\bar{x}^\mu/\partial x^\alpha$ is nonzero and nonsingular everywhere. A trajectory $x^\mu(\tau)$ in the old coordinates becomes $\bar{x}^\mu(x(\tau)) \equiv \bar{x}^\mu(\tau)$ in the new ones, where τ labels a fixed point on the trajectory independently of the coordinates.

The action depends on the metric tensor, one-form potential and velocity components, which under a coordinate transformation change to

$$g_{\bar{\mu}\bar{\nu}} = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial \bar{x}^\mu}\frac{\partial x^\beta}{\partial \bar{x}^\nu} \ , \quad A_{\bar{\mu}} = A_\alpha \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \ , \quad \frac{d\bar{x}^\mu}{d\tau} = \frac{\partial \bar{x}^\mu}{\partial x^\alpha}\frac{dx^\alpha}{d\tau} \ . \tag{11}$$

We have assumed that $\partial\bar{x}^\mu/\partial x^\alpha$ is invertible. Under coordinate transformations the action does not even change form (only the coordinate labels change), so coordinate transformations alone cannot generate any nondynamical symmetries. However, we will show below that coordinate invariance can generate *dynamical* symmetries which apply only to solutions of the Euler-Lagrange equations.

Under a pushforward, the trajectory $x^\mu(\tau)$ is shifted to a different trajectory with coordinates $y^\mu(\tau)$. After the pushforward, we transform the coordinates to $\bar{x}^\mu(y(\tau))$. Because the pushforward is a one-to-one mapping of the manifold to itself, we are free to choose our coordinate transformation so that $\bar{x} = x$, i.e. $\bar{x}^\mu(y(\tau)) \equiv \bar{x}^\mu(\tau) = x^\mu(\tau)$. In other words, we transform the coordinates so that the new coordinates of the new trajectory are the same as the old coordinates of the old trajectory. The pushforward changes the trajectory; the coordinate transformation covers our tracks.

The combination of pushforward and coordinate transformation is an example of a **diffeomorphism**. A diffeomorphism is a one-to-one mapping between the manifold and itself. In our case, the pushforward and transformation depend on one parameter λ and we have a one-parameter family of diffeomorphisms. After a diffeomorphism, the point P in Figure 2 has the same values of the transformed coordinates as the point P_0 has in the original coordinates: $\bar{x}^\mu(\lambda,\tau) = x^\mu(\tau)$.

Naively, it would seem that a diffeomorphism automatically leaves the action unchanged because the coordinates of the trajectory are unchanged. However, the Lagrangian depends not only on the coordinates of the trajectory; it also depends on tensor components that change according to equation (11). More work will be required before we can tell whether the action is invariant under a diffeomorphism. While a coordinate transformation by itself does not change the action, in general a diffeomorphism, because it involves a pushforward, does. A continuous symmetry occurs when a diffeomorphism does *not* change the action. This is the symmetry we will be studying.

The diffeomorphism is an important operation in general relativity. We therefore digress to consider the diffeomorphism in greater detail before returning to examine its effect on the action.

3.1 Infinitesimal Diffeomorphisms and Lie derivatives

In a diffeomorphism, we shift the point at which a tensor is evaluated by pushing it forward using a vector field and then we transform (pull back) the coordinates so that the shifted point has the same coordinate labels as the old point. Since a diffeomorphism maps a manifold back to itself, under a diffeomorphism a rank (m, n) tensor is mapped to another rank (m, n) tensor. This subsection asks how tensors change under diffeomorphisms.

The pushforward mapping may be symbolically denoted ϕ_λ (following Wald 1984, Appendix C). Thus, a diffeomorphism maps a tensor $\mathsf{T}(P_0)$ at point P_0 to a tensor $\bar{\mathsf{T}}(P) \equiv \phi_\lambda \mathsf{T}(P_0)$ such that the coordinate values are unchanged: $\bar{x}^\mu(P) = x^\mu(P_0)$. (See Fig. 2 for the roles of the points P_0 and P.) The diffeomorphism may be regarded as an active coordinate transformation: under a diffeomorphism the spatial point is changed but the coordinates are not.

We illustrate the diffeomorphism by applying it to the components of the one-form $\tilde{A} = A_\mu \tilde{e}^\mu$ in a coordinate basis:

$$\bar{A}_\mu(P_0) \equiv A_\alpha(P) \frac{\partial x^\alpha}{\partial \bar{x}^\mu}(P) , \quad \text{where } \bar{x}^\mu(P) = x^\mu(P_0) . \tag{12}$$

Starting with A_α at point P_0 with coordinates $x^\mu(P_0)$, we push the coordinates forward to point P, we evaluate A_α there, and then we transform the basis back to the coordinate basis at P with new coordinates $\bar{x}^\mu(P)$.

The diffeomorphism is a continuous, one-parameter family of mappings. Thus, a general diffeomorphism may be obtained from the infinitesimal diffeomorphism with pushforward $y^\mu = x^\mu + \xi^\mu d\lambda$. The corresponding coordinate transformation is (to first order in $d\lambda$)

$$\bar{x}^\mu = x^\mu - \xi^\mu d\lambda \tag{13}$$

so that $\bar{x}^\mu(P) = x^\mu(P_0)$. This yields (in the x^μ coordinate system)

$$\bar{A}_\mu(x) \equiv A_\alpha(x+\xi d\lambda)\frac{\partial x^\alpha}{\partial \bar{x}^\mu} = A_\mu(x) + [\xi^\alpha \partial_\alpha A_\mu(x) + A_\alpha(x)\partial_\mu \xi^\alpha]\, d\lambda + O(d\lambda)^2 \ . \quad (14)$$

We have inverted the Jacobian $\partial\bar{x}^\mu/\partial x^\alpha = \delta^\mu{}_\alpha - \partial_\alpha \xi^\mu d\lambda$ to first order in $d\lambda$, $\partial x^\alpha/\partial\bar{x}^\mu = \delta^\alpha{}_\mu + \partial_\mu \xi^\alpha d\lambda + O(d\lambda)^2$. In a similar manner, the infinitesimal diffeomorphism of the metric gives

$$\begin{aligned} \bar{g}_{\mu\nu}(x) &\equiv g_{\alpha\beta}(x+\xi d\lambda)\frac{\partial x^\alpha}{\partial \bar{x}^\mu}\frac{\partial x^\beta}{\partial \bar{x}^\nu} \\ &= g_{\mu\nu}(x) + [\xi^\alpha \partial_\alpha g_{\mu\nu}(x) + g_{\alpha\nu}(x)\partial_\mu \xi^\alpha + g_{\mu\alpha}(x)\partial_\nu \xi^\alpha]\, d\lambda + O(d\lambda)^2 \ . \end{aligned} \quad (15)$$

In general, the infinitesimal diffeomorphism $\bar{\mathsf{T}} \equiv \phi_{\Delta\lambda}\mathsf{T}$ changes the tensor by an amount first-order in $\Delta\lambda$ and linear in $\vec{\xi}$. This change allows us to define a linear operator called the **Lie derivative**:

$$\mathcal{L}_\xi \mathsf{T} \equiv \lim_{\Delta\lambda\to 0} \frac{\phi_{\Delta\lambda}\mathsf{T}(x) - \mathsf{T}(x)}{\Delta\lambda} \quad \text{with } \bar{x}^\mu(P) = x^\mu(P_0) = x^\mu(P) - \xi^\mu \Delta\lambda + O(\Delta\lambda)^2 \ . \quad (16)$$

The Lie derivatives of $A_\mu(x)$ and $g_{\mu\nu}(x)$ follow from equations (14)–(16):

$$\mathcal{L}_\xi A_\mu(x) = \xi^\alpha \partial_\alpha A_\mu + A_\alpha \partial_\mu \xi^\alpha \ , \quad \mathcal{L}_\xi g_{\mu\nu}(x) = \xi^\alpha \partial_\alpha g_{\mu\nu} + g_{\alpha\nu}\partial_\mu \xi^\alpha + g_{\mu\alpha}\partial_\nu \xi^\alpha \ . \quad (17)$$

The first term of the Lie derivative, $\xi^\alpha \partial_\alpha$, corresponds to the pushforward, shifting a tensor to another point in the manifold. The remaining terms arise from the coordinate transformation back to the original coordinate values. As we will show in the next subsection, this combination of terms makes the Lie derivative a tensor in the tangent space at x^μ.

Under a diffeomorphism the transformed tensor components, regarded as functions of coordinates, are evaluated at exactly the same numerical values of the transformed coordinate fields (but a different point in spacetime!) as the original tensor components in the original coordinates. This point is fundamental to the diffeomorphism and therefore to the Lie derivative, and distinguishes the latter from a directional derivative. Thinking of the tensor components as a set of functions of coordinates, we are performing an active transformation: the tensor component functions are changed but they are evaluated at the original values of the coordinates. The Lie derivative generates an infinitesimal diffeomorphism. That is, under a diffeomorphism with pushforward $x^\mu \to x^\mu + \xi^\mu d\lambda$, any tensor T is transformed to $\mathsf{T} + \mathcal{L}_\xi \mathsf{T} d\lambda$.

The fact that the coordinate values do not change, while the tensor fields do, distinguishes the diffeomorphism from a simple coordinate transformation. An important implication is that, in integrals over spacetime volume, the volume element d^4x does not change under a diffeomorphism, while it does change under a coordinate transformation. By contrast, the volume element $\sqrt{-g}\, d^4x$ is invariant under a coordinate transformation but not under a diffeomorphism.

3.2 Properties of the Lie Derivative

The Lie derivative $\mathcal{L}_\xi$ is similar to the directional derivative operator ∇_ξ in its properties but not in its value, except for a scalar where $\mathcal{L}_\xi f = \nabla_\xi f = \xi^\mu \partial_\mu f$. The Lie derivative of a tensor is a tensor of the same rank. To show that it is a tensor, we rewrite the partial derivatives in equation (17) in terms of covariant derivatives in a coordinate basis using the Christoffel connection coefficients to obtain

$$\begin{aligned} \mathcal{L}_\xi A_\mu &= \xi^\alpha \nabla_\alpha A_\mu + A_\alpha \nabla_\mu \xi^\alpha + T^\alpha{}_{\mu\beta} A_\alpha \xi^\beta \ , \\ \mathcal{L}_\xi g_{\mu\nu} &= \xi^\alpha \nabla_\alpha g_{\mu\nu} + g_{\alpha\nu} \nabla_\mu \xi^\alpha + g_{\mu\alpha} \nabla_\nu \xi^\alpha + T^\alpha{}_{\mu\beta} g_{\alpha\nu} \xi^\beta + T^\alpha{}_{\nu\beta} g_{\mu\alpha} \xi^\beta \ , \end{aligned} \tag{18}$$

where $T^\alpha{}_{\mu\beta}$ is the torsion tensor, defined by $T^\alpha{}_{\mu\beta} = \Gamma^\alpha{}_{\mu\beta} - \Gamma^\alpha{}_{\beta\mu}$ in a coordinate basis. The torsion vanishes by assumption in general relativity. Equations (18) show that $\mathcal{L}_\xi A_\mu$ and $\mathcal{L}_\xi g_{\mu\nu}$ are tensors.

The Lie derivative $\mathcal{L}_\xi$ differs from the directional derivative ∇_ξ in two ways. First, the Lie derivative requires no connection: equation (17) gave the Lie derivative solely in terms of partial derivatives of tensor components. [The derivatives of the metric should not be regarded here as arising from the connection; the Lie derivative of any rank $(0,2)$ tensor has the same form as $\mathcal{L}_\xi g_{\mu\nu}$ in eq. 17.] Second, the Lie derivative involves the derivatives of the vector field $\vec{\xi}$ while the covariant derivative does not. The Lie derivative trades partial derivatives of the metric (present in the connection for the covariant derivative) for partial derivatives of the vector field. The directional derivative tells how a fixed tensor field changes as one moves through it in direction $\vec{\xi}$. The Lie derivative tells how a tensor field changes as it is pushed forward along the integral curves of $\vec{\xi}$.

More understanding of the Lie derivative comes from examining the first-order change in a vector expanded in a coordinate basis under a displacement $\vec{\xi} d\lambda$:

$$d\vec{A} = \vec{A}(x + \xi d\lambda) - \vec{A}(x) = A^\mu(x + \xi d\lambda) \vec{e}_\mu(x + \xi d\lambda) - A^\mu(x) \vec{e}_\mu(x) \ . \tag{19}$$

The nature of the derivative depends on how we obtain $\vec{e}_\mu(x + \xi d\lambda)$ from $\vec{e}_\mu(x)$. For the directional derivative ∇_ξ, the basis vectors at different points are related by the connection:

$$\vec{e}_\mu(x + \xi\lambda) = \left(\delta^\beta{}_\mu + d\lambda\, \xi^\alpha \Gamma^\beta{}_{\mu\alpha} \right) \vec{e}_\beta(x) \ \text{ for } \nabla_\xi \ . \tag{20}$$

For the Lie derivative $\mathcal{L}_\xi$, the basis vector is mapped back to the starting point with

$$\vec{e}_\mu(x + \xi d\lambda) = \frac{\partial \bar{x}^\beta}{\partial x^\mu} \vec{e}_\beta(x) = \left(\delta^\beta{}_\mu - d\lambda\, \partial_\mu \xi^\beta \right) \vec{e}_\beta(x) \ \text{ for } \mathcal{L}_\xi \ . \tag{21}$$

Similarly, the basis one-form is mapped using

$$\tilde{e}^\mu(x + \xi d\lambda) = \frac{\partial x^\mu}{\partial \bar{x}^\beta} \tilde{e}^\beta(x) = \left(\delta^\mu{}_\beta + d\lambda\, \partial_\beta \xi^\mu \right) \tilde{e}^\beta(x) \ \text{ for } \mathcal{L}_\xi \ . \tag{22}$$

These mappings ensure that $d\vec{A}/d\lambda = \mathcal{L}_\xi \vec{A}$ is a tangent vector on the manifold.

The Lie derivative of any tensor may be obtained using the following rules: (1) The Lie derivative of a scalar field is the directional derivative, $\mathcal{L}_\xi f = \xi^\alpha \partial_\alpha f = \nabla_\xi f$. (2) The Lie derivative obeys the Liebnitz rule, $\mathcal{L}_\xi(TU) = (\mathcal{L}_\xi T)U + T(\mathcal{L}_\xi U)$, where T and U may be tensors of any rank, with a tensor product or contraction between them. The Lie derivative commutes with contractions. (3) The Lie derivatives of the basis vectors are $\mathcal{L}_\xi \vec{e}_\mu = -\vec{e}_\alpha \partial_\mu \xi^\alpha$. (4) The Lie derivatives of the basis one-forms are $\mathcal{L}_\xi \tilde{e}^\mu = \tilde{e}^\alpha \partial_\alpha \xi^\mu$.

These rules ensure that the Lie derivative of a tensor is a tensor. Using them, the Lie derivative of any tensor may be obtained by expanding the tensor in a basis, e.g. for a rank $(1,2)$ tensor,

$$\begin{aligned} \mathcal{L}_\xi \mathsf{S} &= \mathcal{L}_\xi(S^\mu{}_{\nu\kappa}\vec{e}_\mu \otimes \tilde{e}^\nu \otimes \tilde{e}^\kappa) \equiv (\mathcal{L}_\xi S^\mu{}_{\nu\kappa})\, \vec{e}_\mu \otimes \tilde{e}^\nu \otimes \tilde{e}^\kappa \\ &= [\xi^\alpha \partial_\alpha S^\mu{}_{\nu\kappa} - S^\alpha{}_{\nu\kappa}\partial_\alpha \xi^\mu + S^\mu{}_{\alpha\kappa}\partial_\nu \xi^\alpha + S^\mu{}_{\nu\alpha}\partial_\kappa \xi^\alpha]\, \vec{e}_\mu \otimes \tilde{e}^\nu \otimes \tilde{e}^\kappa \ . \end{aligned} \tag{23}$$

The partial derivatives can be changed to covariant derivatives without change (with vanishing torsion, the connection coefficients so introduced will cancel each other), confirming that the Lie derivative of a tensor really is a tensor.

The Lie derivative of a vector field is an antisymmetric object known also as the commutator or Lie bracket:

$$\mathcal{L}_V \vec{U} = (V^\mu \partial_\mu U^\nu - U^\mu \partial_\mu V^\nu)\vec{e}_\nu \equiv [\vec{V}, \vec{U}] \ . \tag{24}$$

The commutator was introduced in the notes *Tensor Calculus, Part 2*, Section 2.2. With vanishing torsion, $[\vec{V}, \vec{U}] = \nabla_V \vec{U} - \nabla_U \vec{V}$. Using rule (4) of the Lie derivative given after equation (22), it follows at once that the commutator of any pair of coordinate basis vector fields vanishes: $[\vec{e}_\mu, \vec{e}_\nu] = 0$.

3.3 Diffeomorphism-invariance and Killing Vectors

Having defined and investigated the properties of diffeomorphisms and the Lie derivative, we return to the question posed at the beginning of Section 3: How can we tell when the action is translationally invariant? Equation (10) gives the change in the action under a generalized translation or pushforward by the vector field $\vec{\xi}$. However, it is not yet in a form that highlights the key role played by diffeomorphisms.

To uncover the diffeomorphism we must perform the infinitesimal coordinate transformation given by equation (13). To first order in $d\lambda$ this has no effect on the $d\lambda$ term already on the right-hand side of equation (10) but it does add a piece to the unperturbed action. Using equation (11) and the fact that the Lagrangian is a scalar, to $O(d\lambda)$ we obtain

$$\begin{aligned} S[x(\tau)] = \int_{\tau_1}^{\tau_2} L(g_{\mu\nu}, A_\mu, \dot{x}^\mu)\, d\tau = \int_{\tau_1}^{\tau_2} L\left(g_{\alpha\beta}\frac{\partial x^\alpha}{\partial \bar{x}^\mu}\frac{\partial x^\beta}{\partial \bar{x}^\nu}, A_\alpha \frac{\partial x^\alpha}{\partial \bar{x}^\mu}, \frac{dx^\alpha}{d\tau}\frac{\partial \bar{x}^\mu}{\partial x^\alpha}\right) d\tau \\ = S[x(\tau)] + d\lambda \int_{\tau_1}^{\tau_2} \left[\frac{\partial L}{\partial g_{\mu\nu}}(g_{\alpha\nu}\partial_\mu \xi^\alpha + g_{\mu\alpha}\partial_\nu \xi^\alpha) + \frac{\partial L}{\partial A_\mu}(A_\alpha \partial_\mu \xi^\alpha) - \frac{\partial L}{\partial \dot{x}^\mu}\frac{d\xi^\mu}{d\tau}\right] d\tau \ . \end{aligned} \tag{25}$$

The integral multiplying $d\lambda$ always has the value zero for any trajectory $x^\mu(\tau)$ and vector field $\vec{\xi}$ because of the coordinate-invariance of the action. However, it is a special kind of zero because, when added to the pushforward term of equation (10), it gives a diffeomorphism:

$$S[x(\tau)+\xi(x(\tau))d\lambda] = S[x(\tau)] + d\lambda \int_{\tau_1}^{\tau_2} \left[\frac{\partial L}{\partial g_{\mu\nu}} \mathcal{L}_\xi g_{\mu\nu} + \frac{\partial L}{\partial A_\mu} \mathcal{L}_\xi A_\mu \right] d\tau \ . \qquad (26)$$

If the action contains additional fields, under a diffeomorphism we obtain a Lie derivative term for each field.

Thus, we have answered the question of translation-invariance: the action is translationally invariant if and only if the Lie derivative of each tensor field appearing in the Lagrangian vanishes. The uniform translations of Newtonian mechanics are generalized to diffeomorphisms, which include translations, rotations, boosts, and any continuous, one-to-one mapping of the manifold back to itself.

In Newtonian mechanics, translation-invariance leads to a conserved momentum. What about diffeomorphism-invariance? Does it also lead to a conservation law?

Let us suppose that the original trajectory $x^\mu(\tau)$ satisfies the equations of motion before being pushed forward, i.e. the action, with Lagrangian $L(g_{\mu\nu}(x), A_\mu, \dot{x}^\mu)$, is stationary under first-order variations $x^\mu \to x^\mu + \delta x^\mu(x)$ with fixed endpoints $\delta x^\mu(\tau_1) = \delta x^\mu(\tau_2) = 0$. From equation (26) it follows that the action for the shifted trajectory is also stationary, if and only if $\mathcal{L}_\xi g_{\mu\nu} = 0$ and $\mathcal{L}_\xi A_\mu = 0$. (When the trajectory is varied $x^\mu \to x^\mu + \delta x^\mu$, cross-terms $\xi\delta x$ are regarded as being second-order and are ignored.)

If there exists a vector field $\vec{\xi}$ such that $\mathcal{L}_\xi g_{\mu\nu} = 0$ and $\mathcal{L}_\xi A_\mu = 0$, then we can shift solutions of the equations of motion along $\vec{\xi}(x(\tau))$ and generate new solutions. This is a new continuous symmetry called diffeomorphism-invariance, and it generalizes translational-invariance in Newtonian mechanics and special relativity. The result is a dynamical symmetry, which may be deduced by rewriting equation (26):

$$\begin{aligned}
\lim_{\Delta\lambda\to 0} \frac{S[x(\tau)+\xi(x(\tau))\Delta\lambda] - S[x(\tau)]}{\Delta\lambda} &= \int_{\tau_1}^{\tau_2} \left[\frac{\partial L}{\partial g_{\mu\nu}} \mathcal{L}_\xi g_{\mu\nu} + \frac{\partial L}{\partial A_\mu} \mathcal{L}_\xi A_\mu \right] d\tau \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{\partial L}{\partial x^\alpha} \xi^\alpha + \frac{\partial L}{\partial \dot{x}^\mu} \frac{d\xi^\mu}{d\tau} \right] d\tau \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{x}^\mu} \right) \xi^\mu + \frac{\partial L}{\partial \dot{x}^\mu} \frac{d\xi^\mu}{d\tau} \right] d\tau \\
&= \int_{\tau_1}^{\tau_2} \left[\frac{d}{d\tau} (p_\mu \xi^\mu) \right] d\tau \\
&= \left[p_\mu \xi^\mu \right]_{\tau_1}^{\tau_2} \ . \qquad (27)
\end{aligned}$$

All of the steps are straightforward aside from the second line. To obtain this we first expanded the Lie derivatives using equation (17). The terms multiplying ξ^α were then combined to give $\partial L/\partial x^\alpha$ (regarding the Lagrangian as a function of x^μ and $\dot{x}^\mu$). For the terms multiplying the gradient $\partial_\mu \xi^\alpha$, we used $d\xi^\mu(x(\tau))/d\tau = \dot{x}^\alpha \partial_\alpha \xi^\mu$ combined with equation (6) to convert partial derivatives of L with respect to the fields $g_{\mu\nu}$ and A_μ to partial derivatives with respect to $\dot{x}^\mu$. (This conversion is dependent on the Lagrangian, of course, but works for any Lagrangian that is a function of $g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$ and $A_\mu \dot{x}^\mu$.) To obtain the third line we used the assumption that $x^\mu(\tau)$ is a solution of the Euler-Lagrange equations. To obtain the fourth line we used the definition of canonical momentum,

$$p_\mu \equiv \frac{\partial L}{\partial \dot{x}^\mu} \ . \tag{28}$$

For the Lagrangian of equation (6), $p_\mu = g_{\mu\nu}\dot{x}^\nu + qA_\mu$ is not the mechanical momentum (the first term) but also includes a contribution from the electromagnetic field.

Nowhere in equation (27) did we assume that ξ^μ vanishes at the endpoints. The vector field $\vec{\xi}$ is not just a variation used to obtain equations of motion, nor is it a constant; it is an arbitrary small shift.

Theorem: If the Lagrangian is invariant under the diffeomorphism generated by a vector field $\vec{\xi}$, then $\tilde{p}(\vec{\xi}\,) = p_\mu \xi^\mu$ is conserved along curves that extremize the action, i.e. for trajectories obeying the equations of motion.

This result is a generalization of conservation of momentum. The vector field $\vec{\xi}$ may be thought of as the coordinate basis vector field for a cyclic coordinate, i.e. one that does not appear in the Lagrangian. In particular, if $\partial L/\partial x^\alpha = 0$ for a particular coordinate x^α (e.g. $\alpha = 0$), then L is invariant under the diffeomorphism generated by $\vec{e}_\alpha$ so that p_α is conserved.

When gravity is the only force acting on a particle, diffeomorphism-invariance has a purely geometric interpretation in terms of special vector fields known as Killing vectors. Using equation (18) for a manifold with a metric-compatible connection (implying $\nabla_\alpha g_{\mu\nu} = 0$) and vanishing torsion (both of these are true in general relativity), we find that diffeomorphism-invariance implies

$$\mathcal{L}_\xi g_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0 \ . \tag{29}$$

This equation is known as Killing's equation and its solutions are called Killing vector fields, or Killing vectors for short. Thus, our theorem may be restated as follows: If the spacetime has a Killing vector $\vec{\xi}(x)$, then $p_\mu \xi^\mu$ is conserved along any geodesic. A much shorter proof of this theorem follows from $\nabla_V(p_\mu \xi^\mu) = \xi^\mu \nabla_V p_\mu + p_\mu V^\nu \nabla_\nu \xi^\mu$. The first term vanishes by the geodesic equation, while the second term vanishes from Killing's equation with $p^\mu \propto V^\mu$. Despite being longer, however, the proof based on the Lie derivative is valuable because it highlights the role played by a continuous symmetry, diffeomorphism-invariance of the metric.

One is not free to choose Killing vectors; general spacetimes (i.e. ones lacking symmetry) do not have *any* solutions of Killing's equation. As shown in Appendix C.3 of

Wald (1984), a 4-dimensional spacetime has at most 10 Killing vectors. The Minkowski metric has the maximal number, corresponding to the Poincaré group of transformations: three rotations, three boosts, and four translations. Each Killing vector gives a conserved momentum.

The existence of a Killing vector represents a symmetry: the geometry of spacetime as represented by the metric is invariant as one moves in the $\vec{\xi}$-direction. Such a symmetry is known as an isometry. In the perturbation theory view of diffeomorphisms, isometries correspond to perturbations of the coordinates that leave the metric unchanged.

Any vector field can be chosen as one of the coordinate basis fields; the coordinate lines are the integral curves. In Figure 2, the integral curves were parameterized by λ, which becomes the coordinate whose corresponding basis vector is $\vec{e}_\lambda \equiv \vec{\xi}(x)$. For definiteness, let us call this coordinate $\lambda = x^0$. If $\vec{\xi} = \vec{e}_0$ is a Killing vector, then x^0 is a cyclic coordinate and the spacetime is stationary: $\partial_0 g_{\mu\nu} = 0$. In such spacetimes, and only in such spacetimes, p_0 is conserved along geodesics (aside from special cases like the Robertson-Walker spacetimes, where p_0 is conserved for massless but not massive particles because the spacetime is conformally stationary).

Another special feature of spacetimes with Killing vectors is that they have a conserved 4-vector energy-current $S^\nu = \xi_\mu T^{\mu\nu}$. Local stress-energy conservation $\nabla_\mu T^{\mu\nu} = 0$ then implies $\nabla_\nu S^\nu = 0$, which can be integrated over a volume to give the usual form of an integral conservation law. Conversely, spacetimes without Killing vectors do not have an tensor integral energy conservation law, except for spacetimes that are asymptotically flat at infinity. (However, all spacetimes have a conserved energy-momentum pseudotensor, as discussed in the notes *Stress-Energy Pseudotensors and Gravitational Radiation Power*.)

4 Einstein-Hilbert Action for the Metric

We have seen that the action principle is useful not only for concisely expressing the equations of motion; it also enables one to find identities and conservation laws from symmetries of the Lagrangian (invariance of the action under transformations). These methods apply not only to the trajectories of individual particles. They are readily generalized to spacetime fields such as the electromagnetic four-potential A_μ and, most significantly in GR, the metric $g_{\mu\nu}$ itself.

To understand how the action principle works for continuous fields, let us recall how it works for particles. The action is a functional of configuration-space trajectories. Given a set of functions $q^i(t)$, the action assigns a number, the integral of the Lagrangian over the parameter t. For continuous fields the configuration space is a Hilbert space, an infinite-dimensional space of functions. The single parameter t is replaced by the full set of spacetime coordinates. Variation of a configuration-space trajectory, $q^i(t) \to q^i(t) + \delta q^i(t)$, is generalized to variation of the field values at all points of spacetime, e.g.

$g_{\mu\nu}(x) \to g_{\mu\nu}(x) + \delta g_{\mu\nu}(x)$. In both cases, the Lagrangian is chosen so that the action is stationary for trajectories (or field configurations) that satisfy the desired equations of motion. The action principle concisely specifies those equations of motion and facilitates examination of symmetries and conservation laws.

In general relativity, the metric is the fundamental field characterizing the geometric and gravitational properties of spacetime, and so the action must be a functional of $g_{\mu\nu}(x)$. The standard action for the metric is the Hilbert action,

$$S_{\rm G}[g_{\mu\nu}(x)] = \int \frac{1}{16\pi G}\, g^{\mu\nu} R_{\mu\nu} \sqrt{-g}\, d^4x \ . \tag{30}$$

Here, $g = \det g_{\mu\nu}$ and $R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}$ is the Ricci tensor. The factor $\sqrt{-g}$ makes the volume element invariant so that the action is a scalar (invariant under general coordinate transformations). The Einstein-Hilbert action was first shown by the mathematician David Hilbert to yield the Einstein field equations through a variational principle. Hilbert's paper was submitted five days before Einstein's paper presenting his celebrated field equations, although Hilbert did not have the correct field equations until later (for an interesting discussion of the historical issues see L. Corry et al., Science 278, 1270, 1997).

(The Einstein-Hilbert action is a scalar under general coordinate transformations. As we will show in the notes *Stress-Energy Pseudotensors and Gravitational Radiation Power*, it is possible to choose an action that, while not a scalar under general coordinate transformations, still yields the Einstein field equations. The action considered there differs from the Einstein-Hilbert action by a total derivative term. The only real invariance of the action that is required on physical grounds is local Lorentz invariance.)

In the particle actions considered previously, the Lagrangian depended on the generalized coordinates and their first derivatives with respect to the parameter τ. In a spacetime field theory, the single parameter τ is expanded to the four coordinates x^μ. If it is to be a scalar, the Lagrangian for the spacetime metric cannot depend on the first derivatives $\partial_\alpha g_{\mu\nu}$, because $\nabla_\alpha g_{\mu\nu} = 0$ and the first derivatives can all be transformed to zero at a point. Thus, unless one drops the requirement that the action be a scalar under general coordinate transformations, for gravity one is forced to go to second derivatives of the metric. The Ricci scalar $R = g^{\mu\nu} R_{\mu\nu}$ is the simplest scalar that can be formed from the second derivatives of the metric. Amazingly, when the action for matter and all non-gravitational fields is added to the simplest possible scalar action for the metric, the least action principle yields the Einstein field equations.

To look for symmetries of the Einstein-Hilbert action, we consider its change under variation of the functions $g_{\mu\nu}(x)$ with fixed boundary hypersurfaces (the generalization of the fixed endpoints for an ordinary Lagrangian). It proves to be simpler to regard the inverse metric components $g^{\mu\nu}$ as the field variables. The action depends explicitly on $g^{\mu\nu}$ and the Christoffel connection coefficients, $\Gamma^\alpha{}_{\mu\nu}$, the latter appearing in the Ricci tensor in a coordinate basis:

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha{}_{\mu\nu} - \partial_\mu \Gamma^\alpha{}_{\alpha\nu} + \Gamma^\alpha{}_{\mu\nu}\Gamma^\beta{}_{\alpha\beta} - \Gamma^\alpha{}_{\beta\mu}\Gamma^\beta{}_{\alpha\nu} \ . \tag{31}$$

Lengthy algebra shows that first-order variations of $g^{\mu\nu}$ produce the following changes in the quantities appearing in the Einstein-Hilbert action:

$$\begin{aligned}
\delta\sqrt{-g} &= -\frac{1}{2}\sqrt{-g}\,g_{\mu\nu}\delta g^{\mu\nu} = +\frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\delta g_{\mu\nu}\ , \\
\delta\Gamma^{\alpha}{}_{\mu\nu} &= -\frac{1}{2}\left[\nabla_\mu(g_{\nu\lambda}\delta g^{\alpha\lambda}) + \nabla_\nu(g_{\mu\lambda}\delta g^{\alpha\lambda}) - \nabla_\beta(g_{\mu\kappa}g_{\nu\lambda}g^{\alpha\beta}\delta g^{\kappa\lambda})\right]\ , \\
\delta R_{\mu\nu} &= \nabla_\alpha(\delta\Gamma^{\alpha}{}_{\mu\nu}) - \nabla_\mu(\delta\Gamma^{\alpha}{}_{\alpha\nu})\ , \\
g^{\mu\nu}\delta R_{\mu\nu} &= \nabla_\mu\nabla_\nu\left(-\delta g^{\mu\nu} + g^{\mu\nu}g_{\alpha\beta}\delta g^{\alpha\beta}\right)\ , \\
\delta(g^{\mu\nu}R_{\mu\nu}\sqrt{-g}) &= (G_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu})\sqrt{-g}\ ,
\end{aligned} \tag{32}$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ is the Einstein tensor. The covariant derivative ∇_μ appearing in these equations is taken with respect to the zeroth-order metric $g_{\mu\nu}$. Note that, while $\Gamma^{\alpha}{}_{\mu\nu}$ is not a tensor, $\delta\Gamma^{\alpha}{}_{\mu\nu}$ is. Note also that the variations we perform are not necessarily diffeomorphisms (that is, $\delta g_{\mu\nu}$ is not necessarily a Lie derivative), although diffeomorphisms are variations of just the type we are considering (i.e. variations of the tensor component fields for fixed values of their arguments). Equations (32) are straightforward to derive but take several pages of algebra.

Equations (32) give us the change in the gravitational action under variation of the metric:

$$\begin{aligned}
\delta S_{\rm G} &\equiv S_{\rm G}[g^{\mu\nu} + \delta g^{\mu\nu}] - S_{\rm G}[g^{\mu\nu}] \\
&= \frac{1}{16\pi G}\int (G_{\mu\nu}\delta g^{\mu\nu} + \nabla_\mu v^\mu)\sqrt{-g}\,d^4x\ ,\quad v^\mu \equiv \nabla_\nu(-\delta g^{\mu\nu} + g^{\mu\nu}g_{\alpha\beta}\delta g^{\alpha\beta})\ .
\end{aligned} \tag{33}$$

Besides the desired Einstein tensor term, there is a divergence term arising from $g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\mu v^\mu$ which can be integrated using the covariant Gauss' law. This term raises the question of what is fixed in the variation, and what the endpoints of the integration are.

In the action principle for particles (eq. 2), the endpoints of integration are fixed time values, t_1 and t_2. When we integrate over a four-dimensional volume, the endpoints correspond instead to three-dimensional hypersurfaces. The simplest case is when these are hypersurfaces of constant t, in which case the boundary terms are integrals over spatial volume.

In equation (33), the divergence term can be integrated to give the flux of v^μ through the bounding hypersurface. This term involves the derivatives of $\delta g^{\mu\nu}$ normal to the boundary (e.g. the time derivative of $\delta g^{\mu\nu}$, if the endpoints are constant-time hypersurfaces), and is therefore inconvenient because the usual variational principle sets $\delta g^{\mu\nu}$ but not its derivatives to zero at the endpoints. One may either revise the variational principle so that $g^{\mu\nu}$ and $\Gamma^{\alpha}{}_{\mu\nu}$ are independently varied (the Palatini action), or one can add a boundary term to the Einstein-Hilbert action, involving a tensor called the extrinsic curvature, to cancel the $\nabla_\mu v^\mu$ term (Wald, Appendix E.1). In the following we will ignore this term, understanding that it can be eliminated by a more careful treatment.

(The Schrödinger action presented in the later notes *Stress-Energy Pseudotensors and Gravitational Radiation Power* eliminates the $\nabla_\mu v^\mu$ term.)

For convenience below, we introduce a new notation for the integrand of a functional variation, the functional derivative $\delta S/\delta\psi$, defined by

$$\delta S[\psi] \equiv \int \left(\frac{\delta S}{\delta\psi}\right) \delta\psi \sqrt{-g}\, d^4x \ . \tag{34}$$

Here, ψ is any tensor field, e.g. $g^{\mu\nu}$. The functional derivative is strictly defined only when there are no surface terms arising from the variation. Neglecting the surface term in equation (33), we see that $\delta S_{\rm G}/\delta g^{\mu\nu} = (16\pi G)^{-1} G_{\mu\nu}$.

4.1 Stress-Energy Tensor and Einstein Equations

To see how the Einstein equations arise from an action principle, we must add to $S_{\rm G}$ the action for matter, the source of spacetime curvature. Here, "matter" refers to all particles and fields excluding gravity, and specifically includes all the quarks, leptons and gauge bosons in the world (excluding gravitons). At the classical level, one could include electromagnetism and perhaps a simplified model of a fluid. The total action would become a functional of the metric and matter fields. Independent variation of each field yields the equations of motion for that field. Because the metric implicitly appears in the Lagrangian for matter, matter terms will appear in the equation of motion for the metric. This section shows how this all works out for the simplest model for matter, a classical sum of massive particles.

Starting from equation (1), we sum the actions for a discrete set of particles, weighting each by its mass:

$$S_{\rm M} = \sum_a \int -m_a \left(-g_{00} - 2g_{0i}\dot{x}_a^i - g_{ij}\dot{x}_a^i\dot{x}_a^j\right)^{1/2} dt \ . \tag{35}$$

The subscript a labels each particle. We avoid the problem of having no global proper time by parameterizing each particle's trajectory by the coordinate time. Variation of each trajectory, $x_a^i(t) \to x_a^i(t) + \delta x_a^i(t)$ for particle a with $\Delta S_{\rm M} = 0$, yields the geodesic equations of motion.

Now we wish to obtain the equations of motion for the metric itself, which we do by combining the gravitational and matter actions and varying the metric. After a little algebra, equation (33) gives the variation of $S_{\rm G}$; we must add to it the variation of $S_{\rm M}$. Equation (35) gives

$$\delta S_{\rm M} = \int dt \sum_a \frac{1}{2} m_a \frac{V_a^\mu V_a^\nu}{V_a^0} \delta g_{\mu\nu}(x_a^i(t), t) = \int dt \sum_a -\frac{1}{2} m_a \frac{V_{a\mu} V_{a\nu}}{V_a^0} \delta g^{\mu\nu}(x_a^i(t), t) \ . \tag{36}$$

Variation of the metric naturally gives the normalized 4-velocity for each particle, $V_a^\mu = dx^\mu/d\tau_a$ with $V_{a\mu}V_a^\mu = -1$, with a correction factor $1/V_a^0 = d\tau_a/dt$. Now, if we are to combine equations (33) and (36), we must modify the latter to get an integral over 4-volume. This is easily done by inserting a Dirac delta function. The result is

$$\delta S_{\mathrm{M}} = -\int \left[\frac{1}{2}\sum_a \frac{m_a}{\sqrt{-g}} \frac{V_{a\mu}V_{a\nu}}{V_a^0} \delta^3(x^i - x_a^i(t))\right] \delta g^{\mu\nu}(x)\sqrt{-g}\, d^4x \ . \tag{37}$$

The term in brackets may be rewritten in covariant form by inserting an integral over affine parameter with a delta function to cancel it, $\int d\tau_a\, \delta(t - t(\tau_a))(dt/d\tau_a)$. Noting that $V_a^0 = dt/d\tau_a$, we get

$$\delta S_{\mathrm{M}} = -\int \frac{1}{2} T_{\mu\nu}\delta g^{\mu\nu}(x)\sqrt{-g}\, d^4x = +\int \frac{1}{2} T^{\mu\nu}\delta g_{\mu\nu}(x)\sqrt{-g}\, d^4x \ , \tag{38}$$

where the functional differentiation has naturally produced the stress-energy tensor for a gas of particles,

$$T^{\mu\nu} = 2\frac{\delta S_{\mathrm{M}}}{\delta g_{\mu\nu}} = \sum_a \int d\tau_a \frac{\delta^4(x - x(\tau_a))}{\sqrt{-g}}\, m_a V_a^\mu V_a^\nu \ . \tag{39}$$

Aside from the factor $\sqrt{-g}$ needed to correct the Dirac delta function for non-flat coordinates (because $\sqrt{-g}\, d^4x$ is the invariant volume element), equation (39) agrees exactly with the stress-energy tensor worked out in the 8.962 notes *Number-Flux Vector and Stress-Energy Tensor*.

Equation (38) is a general result, and we take it as the definition of the stress-energy tensor for matter (cf. Appendix E.1 of Wald). Thus, given any action S_{M} for particles or fields (matter), we can vary the coordinates or fields to get the equations of motion and vary the metric to get the stress-energy tensor,

$$T^{\mu\nu} \equiv 2\frac{\delta S_{\mathrm{M}}}{\delta g_{\mu\nu}} \ . \tag{40}$$

Taking the action to be the sum of S_{G} and S_{M}, requiring it to be stationary with respect to variations $\delta g^{\mu\nu}$, now gives the Einstein equations:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \ . \tag{41}$$

The pre-factor $(16\pi G)^{-1}$ on S_{G} was chosen to get the correct coefficient in this equation. The matter action is conventionally normalized so that it yields the stress-energy tensor as in equation (38).

4.2 Diffeomorphism Invariance of the Einstein-Hilbert Action

We return to the variation of the Einstein-Hilbert action, equation (33) without the surface term, and consider diffeomorphisms $\delta g^{\mu\nu} = \mathcal{L}_\xi g^{\mu\nu}$:

$$16\pi G\, \delta S_{\mathrm{G}} = \int G_{\mu\nu}(\mathcal{L}_\xi g^{\mu\nu})\sqrt{-g}\, d^4x = -2\int G^{\mu\nu}(\nabla_\mu \xi_\nu)\sqrt{-g}\, d^4x \ . \tag{42}$$

Here, $\vec{\xi}$ is *not* a Killing vector; it is an arbitrary small coordinate displacement. The Lie derivative $\mathcal{L}_\xi g^{\mu\nu}$ has been rewritten in terms of $-\mathcal{L}_\xi g_{\mu\nu}$ using $g^{\mu\alpha}g_{\alpha\nu} = \delta^\mu{}_\nu$. Note that diffeomorphisms are a class of field variations that correspond to mapping the manifold back to itself. Under a diffeomorphism, the integrand of the Einstein-Hilbert action is varied, including the $\sqrt{-g}$ factor. However, as discussed at the end of §3.1, the volume element d^4x is fixed under a diffeomorphism even though it does change under coordinate transformations. The reason for this is apparent in equation (16): under a diffeomorphism, the coordinate values do not change. The pushforward cancels the transformation. If we simply performed either a passive coordinate transformation or pushforward alone, d^4x would not be invariant. Under a diffeomorphism the variation $\delta g_{\mu\nu} = \mathcal{L}_\xi g_{\mu\nu}$ is a tensor on the "unperturbed background" spacetime with metric $g_{\mu\nu}$.

We now show that any scalar integral is invariant under a diffeomorphism that vanishes at the endpoints of integration. Consider the integrand of any action integral, $\Psi\sqrt{-g}$, where Ψ is any scalar constructed out of the tensor fields of the problem; e.g. $\Psi = R/(16\pi G)$ for the Hilbert action. From the first of equations (32) and the Lie derivative of the metric,

$$\mathcal{L}_\xi\sqrt{-g} = \frac{1}{2}\sqrt{-g}\, g^{\mu\nu}\mathcal{L}_\xi g_{\mu\nu} = (\nabla_\alpha\xi^\alpha)\sqrt{-g}\ . \tag{43}$$

Using the fact that the Lie derivative of a scalar is the directional derivative, we obtain

$$\delta S = \int \mathcal{L}_\xi(\Psi\sqrt{-g})\, d^4x = \int (\xi^\mu\nabla_\mu\Psi + \Psi\nabla_\mu\xi^\mu)\sqrt{-g}\, d^4x = \int \Psi\xi^\mu\, d^3\Sigma_\mu\ . \tag{44}$$

We have used the covariant form of Gauss' law, for which $d^3\Sigma_\mu$ is the covariant hypersurface area element for the oriented boundary of the integrated 4-volume. Physically it represents the difference between the spatial volume integrals at the endpoints of integration in time.

For variations with $\xi^\mu = 0$ on the boundaries, $\delta S = 0$. The reason for this is simple: diffeomorphism corresponds exactly to reparameterizing the manifold by shifting and relabeling the coordinates. Just as the action of equation (1) is invariant under arbitrary reparameterization of the path length with fixed endpoints, a spacetime field action is invariant under reparameterization of the coordinates (with no shift on the boundaries). The diffeomorphism differs from a standard coordinate transformation in that the variation is made so that d^4x is invariant rather than $\sqrt{-g}\, d^4x$, but the result is the same: scalar actions are diffeomorphism-invariant.

In considering diffeomorphisms, we do not assume that $g^{\mu\nu}$ extremizes the action. Thus, using $\delta S_{\rm G} = 0$ under diffeomorphisms, we will get an identity rather than a conservation law.

Integrating equation (42) by parts using Gauss's law gives

$$8\pi G\,\delta S_{\rm G} = -\int G^{\mu\nu}\xi_\nu\, d^3\Sigma_\mu + \int \xi_\nu\nabla_\mu G^{\mu\nu}\sqrt{-g}\, d^4x\ . \tag{45}$$

Under reparameterization, the boundary integral vanishes and $\delta S_{\rm G} = 0$ from above, but ξ_ν is arbitrary in the 4-volume integral. Therefore, diffeomorphism-invariance implies

$$\nabla_\mu G^{\mu\nu} = 0 \ . \tag{46}$$

Equation (46) is the famous contracted Bianchi identity. Mathematically, it is an identity akin to equation (4). It may also be regarded as a geometric property of the Riemann tensor arising from the full Bianchi identities,

$$\nabla_\sigma R^\alpha{}_{\beta\mu\nu} + \nabla_\mu R^\alpha{}_{\beta\nu\sigma} + \nabla_\nu R^\alpha{}_{\beta\sigma\mu} = 0 \ . \tag{47}$$

Contracting on α and μ, then multiplying by $g^{\sigma\beta}$ and contracting again gives equation (46). One can also explicitly verify equation (46) using equation (31), noting that $G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu}$ and $R^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}R_{\alpha\beta}$. Wald gives a shorter and more sophisticated proof in his Section 3.2; an even shorter proof can be given using differential forms (Misner et al chapter 15). Our proof, based on diffeomorphism-invariance, is just as rigorous although quite different in spirit from these geometric approaches.

The next step is to inquire whether diffeomorphism-invariance can be used to obtain true conservation laws and not just offer elegant derivations of identities. Before answering this question, we digress to explore an analogous symmetry in electromagnetism.

4.3 Gauge Invariance in Electromagnetism

Maxwell's equations can be obtained from an action principle by adding two more terms to the total action. In SI units these are

$$S_{\rm EM}[A_\mu, g^{\mu\nu}] = \int -\frac{1}{16\pi} F^{\mu\nu}F_{\mu\nu}\sqrt{-g}\, d^4x \ , \quad S_{\rm I}[A_\mu] = \int A_\mu J^\mu \sqrt{-g}\, d^4x \ , \tag{48}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. Note that $g^{\mu\nu}$ is present in $S_{\rm EM}$ implicitly through raising indices of $F_{\mu\nu}$, and that the connection coefficients occurring in $\nabla_\mu A_\nu$ are cancelled in $F_{\mu\nu}$. Electromagnetism adds two pieces to the action, $S_{\rm EM}$ for the free field A_μ and $S_{\rm I}$ for its interaction with a source, the 4-current density J^μ. Previously we considered $S_I = \int qA_\mu \dot{x}^\mu \, d\tau$ for a single particle; now we couple the electromagnetic field to the current density produced by many particles.

The action principle says that the action $S_{\rm EM} + S_{\rm I}$ should be stationary with respect to variations δA_μ that vanish on the boundary. Applying this action principle (left as a homework exercise for the student) yields the equations of motion

$$\nabla_\nu F^{\mu\nu} = 4\pi J^\mu \ . \tag{49}$$

In the language of these notes, the other pair of Maxwell equations, $\nabla_{[\alpha}F_{\mu\nu]} = 0$, arises from a non-dynamical symmetry, the invariance of $S_{\rm EM}[A_\mu]$ under a gauge transformation

$A_\mu \to A_\mu + \nabla_\mu \Phi$. (Expressed using differential forms, $\mathbf{dF} = 0$ because $\mathbf{F} = \mathbf{dA}$ is a closed 2-form. A gauge transformation adds to $\mathbf{F}$ the term $\mathbf{dd}\Phi$, which vanishes for the same reason. See the 8.962 notes *Hamiltonian Dynamics of Particle Motion.*) The source-free Maxwell equations are simple identities in that $\nabla_{[\alpha} F_{\mu\nu]} = 0$ for *any* differentiable A_μ, whether or not it extremizes any action.

If we require the complete action to be gauge-invariant, a new conservation law appears, charge conservation. Under a gauge transformation, the interaction term changes by

$$\begin{aligned} \delta S_{\mathrm{I}} &\equiv S_{\mathrm{I}}[A_\mu + \nabla_\mu \Phi] - S_{\mathrm{I}}[A_\mu] = \int J^\mu (\nabla_\mu \Phi) \sqrt{-g}\, d^4x \\ &= \int \Phi J^\mu \, d^3\Sigma_\mu - \int \Phi (\nabla_\mu J^\mu) \sqrt{-g}\, d^4x \ . \end{aligned} \quad (50)$$

For gauge transformations that vanish on the boundary, gauge-invariance is equivalent to conservation of charge, $\nabla_\mu J^\mu = 0$. This is an example of Noether's theorem: a continuous symmetry generates a conserved current. Gauge invariance is a dynamical symmetry because the action is extremized if and only if J^μ obeys the equations of motion for whatever charges produce the current. (There will be other action terms, such as eq. 35, to give the charges' equations of motion.) Adding a gauge transformation to a solution of the Maxwell equations yields another solution. All solutions necessarily conserve total charge.

Taking a broad view, physicists regard gauge-invariance as a fundamental symmetry of nature, from which charge conservation follows. A similar phenomenon occurs with the gravitational equivalent of gauge invariance, as we discuss next.

4.4 Energy-Momentum Conservation from Gauge Invariance

The example of electromagnetism sheds light on diffeomorphism-invariance in general relativity. We have already seen that every piece of the action is automatically diffeomorphism-invariant because of parameterization-invariance. However, we wish to single out gravity — specifically, the metric $g_{\mu\nu}$ — to impose a symmetry requirement akin to electromagnetic gauge-invariance.

We do this by defining a gauge transformation of the metric as an infinitesimal diffeomorphism,

$$g_{\mu\nu} \to g_{\mu\nu} + \mathcal{L}_\xi g_{\mu\nu} = g_{\mu\nu} + \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu \quad (51)$$

where $\xi^\mu = 0$ on the boundary of our volume. (If the manifold is compact, it has a natural boundary; otherwise we integrate over a compact subvolume. See Appendix A of Wald for mathematical rigor.) Gauge-invariance (diffeomorphism-invariance) of the Einstein-Hilbert action leads to a mathematical identity, the twice-contracted Bianchi identity, equation (46). The rest of the action, including all particles and fields, must also be diffeomorphism-invariant. In particular, this means that the matter action must be invariant under the gauge transformation of equation (51). Using equation (38), this requirement leads to a conservation law:

$$\delta S_{\mathrm{M}} = \int T^{\mu\nu}(\nabla_\mu \xi_\nu)\sqrt{-g}\, d^4x = -\int \xi_\nu (\nabla_\mu T^{\mu\nu})\sqrt{-g}\, d^4x = 0 \quad \Rightarrow \quad \nabla_\mu T^{\mu\nu} = 0 \ . \qquad (52)$$

In general relativity, total stress-energy conservation is a consequence of gauge-invariance as defined by equation (51). Local energy-momentum conservation therefore follows as an application of Noether's theorem (a continuous symmetry of the action leads to a conserved current) just as electromagnetic gauge invariance implies charge conservation.

There is a further analogy with electromagnetism. Physical observables in general relativity must be gauge-invariant. If we wish to try to deduce physics from the metric or other tensors, we will have to work with gauge-invariant quantities or impose gauge conditions to fix the coordinates and remove the gauge freedom. This issue will arise later in the study of gravitational radiation.

5 An Example of Gauge Invariance and Diffeomorphism Invariance: The Ginzburg-Landau Model

The discussion of gauge invariance in the preceding section is incomplete (although fully correct) because under a diffeomorphism all fields change, not only the metric. Similarly, the matter fields for charged particles also change under an electromagnetic gauge transformation and under the more complicated symmetry transformations of non-Abelian gauge symmetries such as those present in the theories of the electroweak and strong interactions. In order to give a more complete picture of the role of gauge symmetries in both electromagnetism and gravity, we present here the classical field theory for the simplest charged field, a complex scalar field $\phi(x)$ representing spinless particles of charge q and mass m. Although there are no fundamental particles with spin 0 and nonzero electric charge, this example is very important in physics as it describes the effective field theory for superconductivity developed by Ginzburg and Landau.

The Ginzburg-Landau model illustrates the essential features of gauge symmetry arising in the standard model of particle physics and its classical extension to gravity. At the classical level, the Ginzburg-Landau model describes a charged fluid, e.g. a fluid of Cooper pairs (the electron pairs that are responsible for superconductivity). Here we couple the charged fluid to gravity as well as to the electromagnetic field.

The Ginzburg-Landau action is (with a sign difference in the kinetic term compared with quantum field theory textbooks because of our choice of metric signature)

$$S_{\mathrm{GL}}[\phi, A_\mu, g^{\mu\nu}] = \int \left[-\frac{1}{2} g^{\mu\nu}(D_\mu\phi)^*(D_\nu\phi) + \frac{1}{2}\mu^2\phi^*\phi - \frac{\lambda}{4}(\phi^*\phi)^2 \right] \sqrt{-g}\, d^4x \ , \qquad (53)$$

where ϕ^* is the complex conjugate of ϕ and

$$D_\mu \equiv \nabla_\mu - iqA_\mu(x) \tag{54}$$

is called the **gauge covariant derivative**. The electromagnetic one-form potential appears so that the action is automatically gauge-invariant. Under an electromagnetic gauge transformation, both the electromagnetic potential and the scalar field change, as follows:

$$A_\mu(x) \to A_\mu(x) + \nabla_\mu\Phi(x) \ , \quad \phi(x) \to e^{iq\Phi(x)}\phi(x) \ , \quad D_\mu\phi \to e^{iq\Phi(x)}D_\mu\phi \ , \tag{55}$$

where $\Phi(x)$ is any real scalar field. We see that $(D^\mu\phi)^*(D_\nu\phi)$ and the Ginzburg-Landau action are gauge-invariant. Thus, an electromagnetic gauge transformation corresponds to an independent change of phase at each point in spacetime, or a local $U(1)$ symmetry.

The gauge covariant derivative automatically couples our charged scalar field to the electromagnetic field so that no explicit interaction term is needed, unlike in equation (48). The first term in the Ginzburg-Landau action is a "kinetic" part that is quadratic in the derivatives of the field. The remaining parts are "potential" terms. The quartic term with coefficient $\lambda/4$ represents the effect of self-interactions that lead to a phenomenon called spontaneous symmetry breaking. Although spontaneous symmetry breaking is of major importance in modern physics, and is an essential feature of the Ginzburg-Landau model, it has no effect on our discussion of symmetries and conservation laws so we ignore it in the following.

The appearance of A_μ in the gauge covariant derivative is reminiscent of the appearance of the connection $\Gamma^\mu_{\alpha\beta}$ in the covariant derivative of general relativity. However, the gravitational connection is absent for derivatives of scalar fields. We will not discuss the field theory of charged vector fields (which represent spin-1 particles in non-Abelian theories) or spinors (spin-1/2 particles).

A complete model includes the actions for gravity and the electromagnetic field in addition to $S_{\rm GL}$: $S[\phi, A_\mu, g^{\mu\nu}] = S_{\rm GL}[\phi, A_\mu, g^{\mu\nu}] + S_{\rm EM}[A_\mu, g^{\mu\nu}] + S_{\rm G}[g^{\mu\nu}]$. According to the action principle, the classical equations of motion follow by requiring the total action to be stationary with respect to small independent variations of $(\phi, A_\mu, g^{\mu\nu})$ at each point in spacetime. Varying the action yields

$$\begin{aligned}
\frac{\delta S}{\delta\phi} &= g^{\mu\nu}D_\mu D_\nu\phi + \left(\mu^2 - \lambda\phi^*\phi\right)\phi \ , \\
\frac{\delta S}{\delta A_\mu} &= -\frac{1}{4\pi}\nabla_\nu F^{\mu\nu} + J^\mu_{\rm GL} \ , \\
\frac{\delta S}{\delta g^{\mu\nu}} &= \frac{1}{16\pi G}G_{\mu\nu} - \frac{1}{2}T^{\rm EM}_{\mu\nu} - \frac{1}{2}T^{\rm GL}_{\mu\nu} \ ,
\end{aligned} \tag{56}$$

where the current and stress-energy tensor of the charged fluid are

$$J^{\rm GL}_\mu \equiv \frac{iq}{2}\left[\phi(D_\mu\phi)^* - \phi^*(D_\mu\phi)\right] \ ,$$

$$T^{\rm GL}_{\mu\nu} \equiv (D_\mu\phi)^*(D_\nu\phi) + \left[-\frac{1}{2}g^{\alpha\beta}(D_\alpha\phi)^*(D_\beta\phi) + \frac{1}{2}\mu^2\phi^*\phi - \frac{\lambda}{4}(\phi^*\phi)^2\right] g_{\mu\nu} \ . \quad (57)$$

The expression for the current density is very similar to the probability current density in nonrelativistic quantum mechanics. The expression for the stress-energy tensor seems strange, so let us examine the energy density in locally Minkowski coordinates (where $g_{\mu\nu} = \eta_{\mu\nu}$):

$$\rho_{\rm GL} = T^{\rm GL}_{00} = \frac{1}{2}\left|D_0\phi\right|^2 + \frac{1}{2}\left|D_i\phi\right|^2 - \frac{1}{2}\mu^2\phi^*\phi + \frac{\lambda}{4}(\phi^*\phi)^2 \ . \quad (58)$$

Aside from the electromagnetic contribution to the gauge covariant derivatives and the potential terms involving $\phi^*\phi$, this looks just like the energy density of a field of relativistic harmonic oscillators. (The potential energy is minimized for $|\phi| = \mu/\sqrt{\lambda}$. This is a circle in the complex ϕ plane, leading to spontaneous symmetry breaking as the field acquires a phase. Those with a knowledge of field theory will recognize two modes for small excitations: a massive mode with mass $\sqrt{2}\mu$ and a massless Goldstone mode corresponding to the field circulating along the circle of minima.)

The equations of motion follow immediately from setting the functional derivatives to zero. The equations of motion for $g^{\mu\nu}$ and A_μ are familiar from before; they are simply the Einstein and Maxwell equations with source including the current and stress-energy of the charged fluid. The equation of motion for ϕ is a nonlinear relativistic wave equation. If $A_\mu = 0$, $\mu^2 = -m^2$, $\lambda = 0$, and $g_{\mu\nu} = \eta_{\mu\nu}$ then it reduces to the Klein-Gordon equation, $(\partial_t^2 - \partial^2 + m^2)\phi = 0$ where $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$ is the spatial Laplacian. Our equation of motion for ϕ generalizes the Klein-Gordon equation to include the effects of gravity (through $g^{\mu\nu}$), electromagnetism (through A_μ), and self-interactions (through $\lambda\phi^*\phi$).

Now we can ask about the consequences of gauge invariance. First, the Ginzburg-Landau current and stress-energy tensor are gauge-invariant, as is easily verified using equations (55) and (57). The action is explicitly gauge-invariant. Using equations (56), we can ask about the effect of an infinitesimal gauge transformation, for which $\delta\phi = iq\Phi(x)\phi$, $\delta A_\mu = \nabla_\mu\Phi$, and $\delta g^{\mu\nu} = 0$. The change in the action is

$$\begin{aligned} \delta S &= \int \left[\frac{\delta S}{\delta\phi}(iq\Phi\phi) + \frac{\delta S}{\delta A_\mu}(\nabla_\mu\Phi)\right]\sqrt{-g}\, d^4x \\ &= \int \left[iq\phi\frac{\delta S}{\delta\phi} - \nabla_\mu\left(\frac{\delta S}{\delta A_\mu}\right)\right]\Phi(x)\sqrt{-g}\, d^4x \ , \end{aligned} \quad (59)$$

where we have integrated by parts and dropped a surface term assuming that $\Phi(x)$ vanishes on the boundary. Now, requiring $\delta S = 0$ under a gauge transformation for the total action adds nothing new because we already required $\delta S/\delta\phi = 0$ and $\delta S/\delta A_\mu = 0$. However, we have constructed each piece of the action ($S_{\rm GL}$, $S_{\rm EM}$ and $S_{\rm G}$) to be gauge-invariant. This gives:

$$\delta S_{\rm GL} = 0 \qquad \Rightarrow \quad iq\phi\frac{\delta S}{\delta\phi} - \nabla_\mu J^\mu_{\rm GL} = 0 \ ,$$
$$\delta S_{\rm EM} = 0 \qquad \Rightarrow \quad -\frac{1}{4\pi}\nabla_\mu\nabla_\nu F^{\mu\nu} = 0 \ . \tag{60}$$

For $S_{\rm GL}$, gauge invariance implies charge conservation provided that the field ϕ obeys the equation of motion $\delta S/\delta\phi = 0$. For $S_{\rm EM}$, gauge invariance gives a trivial identity because $F^{\mu\nu}$ is antisymmetric.

Similar results occur for diffeomorphism invariance, the gravitational counterpart of gauge invariance. Under an infinitesimal diffeomorphism, $\delta\phi = \mathcal{L}_\xi\phi$, $\delta A_\mu = \mathcal{L}_\xi A_\mu$, and $\delta g_{\mu\nu} = \mathcal{L}_\xi g_{\mu\nu} = \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu$. The change in the action is

$$\begin{aligned} \delta S &= \int\left[\frac{\delta S}{\delta\phi}\mathcal{L}_\xi\phi + \frac{\delta S}{\delta A_\mu}\mathcal{L}_\xi A_\mu + \frac{\delta S}{\delta g_{\mu\nu}}\mathcal{L}_\xi g_{\mu\nu}\right]\sqrt{-g}\,d^4x \\ &= \int\left[\frac{\delta S}{\delta\phi}\xi^\mu\nabla_\mu\phi + \left(-\frac{1}{4\pi}\nabla_\nu F^{\mu\nu} + J^\mu\right)\mathcal{L}_\xi A_\mu + \right. \\ &\qquad \left. + \left(-\frac{1}{8\pi G}G^{\mu\nu} + T^{\mu\nu}\right)\nabla_\mu\xi_\nu\right]\sqrt{-g}\,d^4x \ , \end{aligned} \tag{61}$$

where $J^\mu = J^\mu_{\rm GL}$ and $T^{\mu\nu} = T^{\mu\nu}_{\rm GL} + T^{\mu\nu}_{\rm EM}$. As above, requiring that the total action be diffeomorphism-invariant adds nothing new. However, we have constructed each piece of the action to be diffeomorphism-invariant, i.e. a scalar under general coordinate transformations. Applying diffeomorphism-invariance to $S_{\rm GL}$ gives a subset of the terms in equation (61),

$$\begin{aligned} 0 &= \int\left[\frac{\delta S}{\delta\phi}\xi^\mu\nabla_\mu\phi + J^\mu\left(\xi^\alpha\nabla_\alpha A_\mu + A_\alpha\nabla_\mu\xi^\alpha\right) + T^{\mu\nu}_{\rm GL}\nabla_\mu\xi_\nu\right]\sqrt{-g}\,d^4x \\ &= \int\left[-\frac{\delta S}{\delta\phi}\nabla_\mu\phi + J^\alpha\nabla_\mu A_\alpha - \nabla_\alpha\left(J^\alpha A_\mu\right) - \nabla^\nu T^{\rm GL}_{\mu\nu}\right]\xi^\mu(x)\sqrt{-g}\,d^4x \\ &= \int\left[-\frac{\delta S}{\delta\phi}\nabla_\mu\phi - (\nabla_\alpha J^\alpha)A_\mu + J^\alpha F_{\mu\alpha} - \nabla^\nu T^{\rm GL}_{\mu\nu}\right]\xi^\mu(x)\sqrt{-g}\,d^4x \ , \end{aligned} \tag{62}$$

where we have discarded surface integrals in the second line assuming that $\xi^\mu(x) = 0$ on the boundary.

Equation (62) gives a nice result. First, as always, our continuous symmetry (here, diffeomorphism-invariance) only gives physical results for solutions of the equations of motion. Thus, $\delta S/\delta\phi = \nabla_\alpha J^\alpha = 0$ can be dropped without further consideration. The remaining terms individually need not vanish from the equations of motion. From this we conclude

$$\nabla_\nu T^{\mu\nu}_{\rm GL} = F^{\mu\nu}J^{\rm GL}_\nu \ . \tag{63}$$

This has a simple interpretation: the work done by the electromagnetic field transfers energy-momentum to the charged fluid. Recall that the Lorentz force on a single charge with 4-velocity V^μ is $qF^{\mu\nu}V_\nu$ and that 4-force is the rate of change of 4-momentum. The current qV^μ for a single charge becomes the current density J^μ of a continuous fluid. Thus, equation (63) gives energy conservation for the charged fluid, including the transfer of energy to and from the electromagnetic field.

The reader can show that requiring $\delta S_{\mathrm{EM}} = 0$ under an infinitesimal diffeomorphism proceeds in a very similar fashion to equation (62) and yields the result

$$\nabla_\mu T^{\mu\nu}_{\mathrm{EM}} = -F^{\mu\nu} J^{\mathrm{GL}}_\nu \ . \tag{64}$$

This result gives the energy-momentum transfer from the viewpoint of the electromagnetic field: work done by the field on the fluid removes energy from the field. Combining equations (63) and (64) gives conservation of total stress-energy, $\nabla_\mu T^{\mu\nu} = 0$.

Finally, because S_{G} depends only on $g^{\mu\nu}$ and not on the other fields, diffeomorphism invariance yields the results already obtained in equations (45) and (46).