

Lecture 7: Introduction to General Relativity

For the past six chapters, we have carefully constructed the mathematical machinery of differential geometry. We defined manifolds to escape the rigid grid of flat Euclidean space, introduced the metric tensor to measure distances, used the covariant derivative to properly transport vectors across curved spaces, and established differential forms to rigorously integrate over these geometries.

Now, we bridge the gap between pure mathematics and physics. We will take the geometric framework we have built and use it to describe gravity not as a force, but as the manifestation of spacetime curvature. This is the heart of Albert Einstein's General Theory of Relativity.

1 Introduction to Special Relativity

Before leaping into curved manifolds, we must firmly establish the physical baseline: the flat spacetime of Special Relativity (SR). Formulated by Albert Einstein in 1905, SR overthrew the Newtonian concepts of absolute space and absolute, universally ticking time, weaving them instead into a single, rigid four-dimensional fabric known as Minkowski spacetime.

1.1 The Postulates of Special Relativity

The entire geometric and kinematic framework of Special Relativity rests strictly upon two deceptively simple physical postulates:

- **The Principle of Relativity:** The laws of physics take the exact same mathematical form (are invariant) in all inertial frames of reference. There is no absolute state of rest, and no experiment can determine whether a non-accelerating frame is moving or stationary.
- **The Constancy of the Speed of Light:** The speed of light in a vacuum, c , is exactly the same for all inertial observers, completely independent of the relative motion of the light source or the observer.

1.2 The Lorentz Transformation

To satisfy the rigid requirement that the speed of light remains constant for everyone, the classical Galilean transformations must be discarded. If a primed frame (t', x', y', z') is moving with a constant velocity v along the positive x -axis relative to an unprimed resting frame (t, x, y, z) , their coordinates are related by the **Lorentz Transformation**:

$$t' = \gamma \left(t - \frac{vx}{c^2} \right) \quad (1)$$

$$x' = \gamma(x - vt) \quad (2)$$

$$y' = y \quad (3)$$

$$z' = z \quad (4)$$

Here, $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the Lorentz factor. These equations demonstrate mathematically that space and time are no longer distinct; they mix and rotate into one another depending on the observer's relative velocity.

1.3 The Invariant Interval and $ds^2 = 0$

In 3D Euclidean space, the absolute distance between two points is an invariant quantity that all observers agree on. In 4D Minkowski spacetime, different observers will disagree on lengths (length contraction) and time intervals (time dilation), but they all universally agree on the **spacetime**

interval, ds^2 . Using the signature $(-, +, +, +)$ and setting $c = 1$, the infinitesimal interval is defined as:

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \quad (5)$$

A profoundly important physical geometric case occurs when $ds^2 = 0$. This describes a **null interval**, which maps the exact trajectory of a photon (a massless particle of light) moving through spacetime. Physically, $ds^2 = 0$ means that the spatial distance traveled perfectly equals the time elapsed ($dx^2 + dy^2 + dz^2 = dt^2$), natively encoding the second postulate into the geometry itself.

Furthermore, the condition $ds^2 = 0$ rigorously defines the **light cone** at any given point in spacetime. It establishes the rigid causal structure of the universe, strictly separating events that can physically influence each other ($ds^2 < 0$, timelike trajectories for massive particles) from those that cannot ($ds^2 > 0$, spacelike intervals).

In Special Relativity, these light cones point rigidly in the exact same direction everywhere across the flat manifold. As we transition to General Relativity, we will discover that the presence of mass and energy curves this manifold, causing these light cones to tilt and converge, fundamentally producing the physical phenomenon we experience as gravity.

2 The Principle of Equivalence and Local Flatness

In Chapter 1, we defined a manifold by its defining intuition: *locally flat, but globally curved*. If you zoom in closely enough on any smooth manifold, it is indistinguishable from flat Euclidean space (or in physics, flat Minkowski spacetime).

Einstein realized that this purely mathematical property perfectly mirrors a profound physical observation: **The Equivalence Principle**. Einstein postulated that in any small, freely falling laboratory (like an elevator whose cable has snapped), the laws of physics behave exactly as they do in the absence of gravity. Locally, gravity “disappears.”

Mathematically, this means that at any specific point p in spacetime, we can always choose a local coordinate chart (a “freely falling frame”) where the metric tensor reduces to the flat Minkowski metric $\eta_{\mu\nu}$, and all the Christoffel symbols (connection coefficients) vanish:

$$g_{\mu\nu}(p) = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Gamma_{\mu\nu}^\alpha(p) = 0 \quad (6)$$

Because $\Gamma_{\mu\nu}^\alpha = 0$ at point p , the covariant derivative ∇_μ temporarily reduces to the standard partial derivative ∂_μ . This provides a powerful heuristic for translating laws from Special Relativity into General Relativity, often called the *Comma-to-Semicolon Rule*: Take any valid tensor equation in flat spacetime and replace the partial derivatives (often denoted by commas) with covariant derivatives (denoted by semicolons).

Concept Check 1: We can choose coordinates to make $g_{\mu\nu} = \eta_{\mu\nu}$ and the first derivatives $\partial_\gamma g_{\mu\nu} = 0$ at a point p . However, we generally *cannot* make the second derivatives $\partial_\gamma \partial_\lambda g_{\mu\nu} = 0$ vanish. Based on our discussion in Chapter 3, what geometric object prevents us from zeroing out these second derivatives, proving that true gravity cannot be globally transformed away?

3 General Coordinate Invariance

In classical physics, we often define our physical laws in highly specific coordinate systems. However, a fundamental tenet of differential geometry is that the underlying manifold exists independently of the coordinates we map onto it.

Einstein elevated this mathematical necessity to a physical law known as **General Covariance** or **Diffeomorphism Invariance**: The fundamental laws of physics must take the exact same mathematical form regardless of the coordinate system used by the observer.

How do we guarantee this? By writing our physical laws exclusively as **tensor equations**. Tensors are geometric objects that exist independently of coordinates. If a tensor equation like $A^\mu = B^\mu$ holds true in one coordinate system, the tensor transformation rules guarantee it will hold true in *all* valid coordinate systems.

For example, the path of a free particle in flat spacetime is a straight line: $\frac{d^2 x^\mu}{d\tau^2} = 0$. To make this universally valid in any curved, gravitating spacetime, we replace the standard coordinate acceleration with the covariant geometric acceleration. This perfectly recovers the Geodesic Equation derived in Chapter 2:

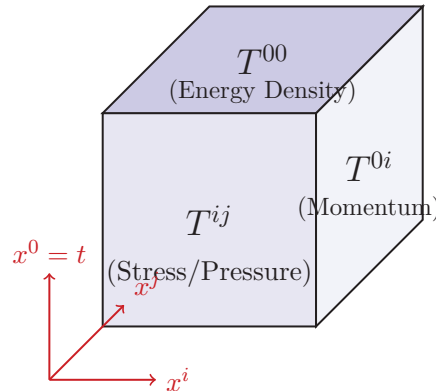
$$u^\nu \nabla_\nu u^\mu = \frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\sigma\nu} \frac{dx^\sigma}{d\tau} \frac{dx^\nu}{d\tau} = 0 \quad (7)$$

Gravity is no longer a force pulling the particle off a straight line; gravity *is* the curvature of the manifold, and the particle is simply tracing the straightest possible path (geodesic) through that curved geometry.

4 The Energy-Momentum Tensor ($T_{\mu\nu}$)

If gravity is the curvature of the spacetime manifold, what causes the manifold to curve? In Newtonian gravity, mass is the source of the gravitational field. In relativity, mass and energy are equivalent ($E = mc^2$). Furthermore, pressure, stress, and momentum also contribute to the energy density of a system.

To describe the distribution of matter and energy in a coordinate-invariant way, we construct a symmetric rank-2 tensor called the **Energy-Momentum Tensor** (or Stress-Energy Tensor), denoted $T^{\mu\nu}$.



The components of $T^{\mu\nu}$ have distinct physical interpretations in a given local coordinate frame:

- T^{00} : The relativistic mass-energy density of the matter.
- T^{0i} and T^{i0} : The momentum density (flux of energy across spatial surfaces).

- T^{ij} : The spatial stress tensor. Diagonal elements (T^{11}, T^{22}, T^{33}) represent pressure, while off-diagonal elements represent shear stress.

4.1 Example: The Perfect Fluid

In astrophysics and cosmology, we frequently model stars or the universe itself as a “perfect fluid”—a fluid with no viscosity or heat conduction. Let ρ be the rest-frame energy density, P be the isotropic pressure, and u^μ be the fluid’s 4-velocity vector. The energy-momentum tensor is elegantly written as:

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + P g^{\mu\nu} \quad (8)$$

If the fluid is static “dust” (zero pressure, $P = 0$), this reduces simply to $T^{\mu\nu} = \rho u^\mu u^\nu$.

5 Covariant Conservation of Energy-Momentum

In flat space, the conservation of energy and momentum is expressed by the continuity equation $\partial_\mu T^{\mu\nu} = 0$. Applying our “comma-to-semicolon” rule to ensure general coordinate invariance, we elevate this to the **Covariant Conservation Law**:

$$\nabla_\mu T^{\mu\nu} = 0 \quad (9)$$

While this looks like a simple generalization, the geometric implications are profound. Expanding the covariant derivative reveals the hidden Christoffel symbols:

$$\partial_\mu T^{\mu\nu} + \Gamma^\mu_{\sigma\mu} T^{\sigma\nu} + \Gamma^\nu_{\sigma\mu} T^{\mu\sigma} = 0 \quad (10)$$

Because the partial derivative $\partial_\mu T^{\mu\nu}$ does not equal zero on its own, *energy-momentum is not globally conserved for the matter alone in a curved spacetime*. The extra terms involving Γ represent the exchange of energy and momentum between the matter field and the gravitational field (the curvature of the manifold itself).

Concept Check 2: In Chapter 4, we learned that continuous symmetries lead to conservation laws via Killing vectors. Suppose the spacetime geometry possesses a continuous symmetry characterized by a Killing vector field K_ν (meaning $\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0$). Define a current vector $J^\mu = T^{\mu\nu} K_\nu$. Using the covariant conservation of $T^{\mu\nu}$ and the properties of the Killing vector, prove that this current is strictly conserved, i.e., show that $\nabla_\mu J^\mu = 0$.

6 Towards the Field Equations

We now have two distinct domains: the geometry of the manifold (characterized by the metric $g_{\mu\nu}$ and the Riemann curvature tensor $R^\rho_{\sigma\mu\nu}$) and the matter content residing within the manifold (characterized by the energy-momentum tensor $T_{\mu\nu}$).

Einstein’s revolutionary insight was that these two domains are dynamically coupled—gravity is not a force propagating through space, but rather the curvature of spacetime itself, dictated by the presence of matter and energy. To formalize this as a tensor equation, we must find a purely geometric tensor $G_{\mu\nu}$, built entirely from the metric and Riemann tensors, that can be set proportional to the energy-momentum tensor:

$$G_{\mu\nu} \propto T_{\mu\nu} \quad (11)$$

Because the energy-momentum tensor satisfies the covariant continuity equation ($\nabla_\mu T^{\mu\nu} = 0$), pure mathematical consistency demands that our chosen geometric tensor *must* also have zero covariant

divergence:

$$\nabla_\mu G^{\mu\nu} = 0 \quad (12)$$

Our goal now is to rigorously derive the exact form of $G_{\mu\nu}$ that satisfies this stringent identity, leading directly to the Einstein Tensor and the field equations.

7 Symmetries of the Riemann Tensor

To construct $G_{\mu\nu}$, we begin with the Riemann curvature tensor, which contains all information about the intrinsic curvature of the manifold. When its first index is lowered using the metric ($R_{\rho\sigma\mu\nu} = g_{\rho\lambda} R^\lambda_{\sigma\mu\nu}$), the tensor exhibits several profound algebraic symmetries:

- **Antisymmetry in the last two indices:** $R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$
- **Antisymmetry in the first two indices:** $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$
- **Symmetry under pair exchange:** $R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$
- **The First Bianchi Identity (cyclic sum):** $R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0$

8 The Second Bianchi Identity

To find a tensor with zero covariant divergence, we must examine the covariant derivatives of the curvature. The relationship between the gradients of curvature across the manifold is given by the **Second Bianchi Identity**:

$$\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\mu R_{\rho\sigma\nu\lambda} + \nabla_\nu R_{\rho\sigma\lambda\mu} = 0 \quad (13)$$

(Geometrically, this is the tensor equivalent of the vector calculus identity stating that the boundary of a boundary is zero, a concept we thoroughly explored in Chapter 6).

9 Contracting the Curvature

The Second Bianchi Identity is a rank-5 tensor equation. To build our rank-2 tensor $G_{\mu\nu}$, we must systematically perform tensor contractions (multiplying by the inverse metric $g^{\alpha\beta}$ to sum over paired indices).

Step A: The Ricci Tensor

First, we contract the first and third indices of the Riemann tensor to define the symmetric **Ricci Tensor**, $R_{\sigma\nu}$:

$$R_{\sigma\nu} = g^{\rho\mu} R_{\rho\sigma\mu\nu} = R^\mu_{\sigma\mu\nu} \quad (14)$$

Step B: The Ricci Scalar

Next, we contract the Ricci tensor with the inverse metric to find its trace. This produces a single, coordinate-independent number at each point in space called the **Ricci Scalar**, R :

$$R = g^{\sigma\nu} R_{\sigma\nu} = R^\nu_{\nu} \quad (15)$$

Step C: Contracting the Bianchi Identity

We now apply the inverse metrics $g^{\rho\mu}$ and $g^{\sigma\nu}$ to contract the Second Bianchi Identity entirely down to rank-2.

First contraction (multiplying by $g^{\rho\mu}$):

$$g^{\rho\mu} (\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\mu R_{\rho\sigma\nu\lambda} + \nabla_\nu R_{\rho\sigma\lambda\mu}) = 0 \quad (16)$$

$$\nabla_\lambda R_{\sigma\nu} + \nabla_\mu R^\mu_{\sigma\nu\lambda} - \nabla_\nu R_{\sigma\lambda} = 0 \quad (17)$$

Second contraction (multiplying by $g^{\sigma\nu}$):

$$g^{\sigma\nu}(\nabla_\lambda R_{\sigma\nu} + \nabla_\mu R^\mu_{\sigma\nu\lambda} - \nabla_\nu R_{\sigma\lambda}) = 0 \quad (18)$$

$$\nabla_\lambda R - \nabla_\mu R^\mu_\lambda - \nabla^\sigma R_{\sigma\lambda} = 0 \quad (19)$$

Because the Ricci tensor is symmetric, we can safely relabel the dummy indices in the last two terms to see that they are identical:

$$\nabla_\lambda R - 2\nabla_\mu R^\mu_\lambda = 0 \quad (20)$$

To put this in the standard contravariant form, we divide the entire equation by -2 and raise the λ index to ν using the metric $g^{\nu\lambda}$:

$$\nabla_\mu R^{\mu\nu} - \frac{1}{2}\nabla^\nu R = 0 \quad (21)$$

Finally, because metric compatibility guarantees that the covariant derivative of the metric is identically zero ($\nabla_\mu g^{\mu\nu} = 0$), we can pull the covariant derivative operator completely outside the expression:

$$\nabla_\mu \left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu} R \right) = 0 \quad (22)$$

10 The Einstein Tensor and the Field Equations

The mathematical expression inside those parentheses is the exact geometric object we have been searching for! We define the **Einstein Tensor**, $G^{\mu\nu}$:

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu} R \quad (23)$$

By construction from the Bianchi identities, it satisfies the required covariant conservation:

$$\nabla_\mu G^{\mu\nu} = 0 \quad (24)$$

Because the Energy-Momentum tensor also guarantees $\nabla_\mu T^{\mu\nu} = 0$, we can now confidently equate the geometry of spacetime to the matter contained within it, up to an unknown proportionality constant κ :

$$G_{\mu\nu} = \kappa T_{\mu\nu} \quad (25)$$

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} R = \kappa T_{\mu\nu} \quad (26)$$

By requiring that this tensor equation seamlessly reproduces Newton's law of universal gravitation ($\nabla^2\Phi = 4\pi G\rho$) in the limit of weak gravitational fields and slow-moving particles, the constant is strictly derived to be $\kappa = \frac{8\pi G}{c^4}$.

11 Derivation via the Principle of Least Action

The geometric derivation using the contracted Bianchi identities establishes mathematical consistency, but the fundamental physical approach relies on the calculus of variations applied to an action functional. We extremize the Einstein-Hilbert action coupled with a matter action.

The total action S of the universe is the sum of the gravitational action S_{EH} and the matter action S_M :

$$S = S_{EH} + S_M = \int \left(\frac{1}{2\kappa} R + \mathcal{L}_M \right) \sqrt{-g} d^4x \quad (27)$$

To find the equations of motion, we apply the calculus of variations and demand that the total action is stationary with respect to arbitrary variations of the inverse metric tensor, meaning $\delta S = 0$.

Taking the variation of the gravitational part yields:

$$\delta S_{EH} = \frac{1}{2\kappa} \int [\delta R \sqrt{-g} + R \delta(\sqrt{-g})] d^4x \quad (28)$$

This requires evaluating two variations: the metric determinant $\delta(\sqrt{-g})$ and the Ricci scalar δR .

1. Variation of the Metric Determinant

Using Jacobi's formula for the derivative of a determinant, $\delta(\det g) = \det(g) g^{\mu\nu} \delta g_{\mu\nu}$. Because the metric contraction yields $g_{\mu\nu} g^{\mu\nu} = \delta_\mu^\mu = 4$, the variation of the inverse metric relates to the metric via $\delta g_{\mu\nu} = -g_{\mu\alpha} g_{\nu\beta} \delta g^{\alpha\beta}$. Applying this to the square root of the negative determinant gives:

$$\delta(\sqrt{-g}) = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu} \quad (29)$$

2. Variation of the Ricci Scalar

The Ricci scalar is the contraction of the Ricci tensor, $R = g^{\mu\nu} R_{\mu\nu}$. Its variation requires the product rule:

$$\delta R = R_{\mu\nu} \delta g^{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} \quad (30)$$

To evaluate $\delta R_{\mu\nu}$, we use the **Palatini identity**. Although the Christoffel symbols $\Gamma_{\mu\nu}^\lambda$ are not tensors, their variations $\delta \Gamma_{\mu\nu}^\lambda$ represent the difference between two connections, which *is* a valid tensor. The variation of the Ricci tensor can be written as the covariant derivative of these tensor variations:

$$\delta R_{\mu\nu} = \nabla_\lambda (\delta \Gamma_{\mu\nu}^\lambda) - \nabla_\nu (\delta \Gamma_{\mu\lambda}^\lambda) \quad (31)$$

Consequently, the term $g^{\mu\nu} \delta R_{\mu\nu}$ becomes a total covariant divergence:

$$g^{\mu\nu} \delta R_{\mu\nu} = \nabla_\sigma \left(g^{\mu\nu} \delta \Gamma_{\mu\nu}^\sigma - g^{\mu\sigma} \delta \Gamma_{\mu\lambda}^\lambda \right) \quad (32)$$

By Stokes' theorem, the integral of a total divergence over the manifold transforms into a boundary term. Because the variations $\delta g^{\mu\nu}$ are defined to vanish at the boundary of the integration region, this entire boundary integral evaluates to strictly zero.

3. Synthesizing the Gravitational Variation

Substituting the non-zero terms back into the variation of the Einstein-Hilbert action:

$$\delta S_{EH} = \frac{1}{2\kappa} \int \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^{\mu\nu} \sqrt{-g} d^4x \quad (33)$$

4. Incorporating the Matter Action

The energy-momentum tensor $T_{\mu\nu}$ is formally defined as the response of the matter Lagrangian to changes in the metric geometry:

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta(\mathcal{L}_M \sqrt{-g})}{\delta g^{\mu\nu}} \quad (34)$$

Rearranging this gives the variation of the matter action:

$$\delta S_M = -\frac{1}{2} \int T_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x \quad (35)$$

5. The Final Equations

Setting the total variation $\delta S_{EH} + \delta S_M = 0$ requires the integrands to sum to zero for any arbitrary metric variation $\delta g^{\mu\nu}$:

$$\int \left[\frac{1}{2\kappa} \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) - \frac{1}{2} T_{\mu\nu} \right] \delta g^{\mu\nu} \sqrt{-g} d^4x = 0 \quad (36)$$

Factoring out the shared integration terms isolates the underlying tensor relationship:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu} \quad (37)$$

This recovers the exact field equations derived previously through the contracted Bianchi identities, achieving the profound result fundamentally through the principle of least action.

12 The Cosmological Constant

Before finalizing the theory, Einstein noticed one lingering mathematical ambiguity. Because the covariant derivative of the metric tensor is exactly zero, we can add any constant multiple of the metric to the Einstein tensor without violating the core conservation law $\nabla_\mu (G^{\mu\nu} + \Lambda g^{\mu\nu}) = 0$. This constant, Λ , is the **Cosmological Constant**.

Including this term yields the complete, final form of **Einstein's Field Equations**:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (38)$$

This elegant equation is the master key to General Relativity. The left side describes how the canvas of the universe curves and twists; the right side defines the massive and energetic actors moving upon it. As John Archibald Wheeler famously summarized: *“Spacetime tells matter how to move; matter tells spacetime how to curve.”*

Concept Check 3: Taking the trace of both sides of Einstein's field equations (by multiplying the entire equation by $g^{\mu\nu}$) and temporarily setting $\Lambda = 0$, express the Ricci scalar R strictly in terms of the trace of the energy-momentum tensor $T = T^\mu_\mu$. Based on this result, what happens to the scalar curvature of spacetime in a region filled purely with electromagnetic radiation (which possesses a traceless energy-momentum tensor, $T = 0$)?

Addendum: Geometric Consistency and Diffeomorphism Invariance

Prior to General Relativity, physicists used the “canonical” stress-energy tensor derived from Noether's theorem by translating flat spacetime coordinates. That canonical tensor is often asymmetric and cumbersome. The Hilbert definition vastly supersedes it by automatically guaranteeing two critical geometric properties required to couple matter to the Einstein tensor:

- **Manifest Symmetry:** Because the metric tensor $g^{\mu\nu}$ is intrinsically symmetric, taking a derivative with respect to it mathematically guarantees that $T_{\mu\nu} = T_{\nu\mu}$. This symmetry naturally ensures the conservation of angular momentum without requiring ad-hoc modifications (like the Belinfante-Rosenfeld procedure).
- **Covariant Conservation:** The matter action S_M is a coordinate-independent scalar, meaning it must be invariant under a general diffeomorphism $x^\mu \rightarrow x^\mu + \xi^\mu$. Under this transformation, the metric changes by the Lie derivative $\delta g^{\mu\nu} = \nabla^\mu \xi^\nu + \nabla^\nu \xi^\mu$. If we insert this into the definition of $T_{\mu\nu}$ and demand $\delta S_M = 0$, integrating by parts immediately yields $\nabla_\mu T^{\mu\nu} = 0$.

Therefore, defining $T_{\mu\nu}$ via the variation of the metric is the only way to generate a source tensor that perfectly mirrors the zero covariant divergence of the Einstein tensor ($\nabla_\mu G^{\mu\nu} = 0$), mathematically satisfying the contracted Bianchi identities.