# Solutions

## Exercise 1 — Timelike geodesics and stationary proper time

### Solution to 1.

Write $\dot{x}^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\lambda$ and

$$L(x,\dot{x}) = \sqrt{-g_{\alpha\beta}(x)\dot{x}^\alpha\dot{x}^\beta}.$$

The overall factor $1/c$ in $\tau = (1/c)\int L\,\mathrm{d}\lambda$ does not affect the stationary curve. The required derivatives are

$$\frac{\partial L}{\partial \dot{x}^\mu} = -\frac{g_{\mu\nu}\dot{x}^\nu}{L}, \qquad \frac{\partial L}{\partial x^\mu} = -\frac{1}{2L}\,\partial_\mu g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta.$$

The endpoint term vanishes because $\delta x^\mu = 0$ at $A$ and $B$. The Euler–Lagrange equation therefore gives

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{g_{\mu\nu}\dot{x}^\nu}{L}\right) - \frac{1}{2L}\partial_\mu g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0.$$

Multiplying by $L$, expanding the derivative, and raising the free index with $g^{\rho\mu}$ yields

$$\ddot{x}^\rho + \Gamma^\rho{}_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \frac{\dot{L}}{L}\dot{x}^\rho.$$

Thus the stationary curve is a geodesic written with a possibly non-affine parameter, with

$$f(\lambda) = \frac{\dot{L}}{L} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\ln L.$$

Along a timelike curve, $L = c\,\mathrm{d}\tau/\mathrm{d}\lambda$. If we choose $\lambda = \tau$, then $L = c$ is constant and $f = 0$. The equation becomes

$$\boxed{\frac{\mathrm{d}^2x^\rho}{\mathrm{d}\tau^2} + \Gamma^\rho{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0.}$$

Any parameter $a\tau + b$, with constants $a \neq 0$ and $b$, is also affine.
The functional is invariant under any monotonic reparametrization because $L\,\mathrm{d}\lambda = c\,\mathrm{d}\tau$ is a geometric scalar. Hence the coordinate calculation describes a coordinate-independent curve. A freely falling massive particle follows this timelike geodesic. Between sufficiently nearby events, before any conjugate point is encountered, the geodesic locally *maximizes* proper time; "stationary" is the more general variational statement.

## Exercise 2 — Super-Hamiltonian formulation

### Solution to 2 (a).

For the quadratic Lagrangian $\mathcal{L} = \frac{1}{2}g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta$,

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = g_{\mu\nu}\dot{x}^\nu, \qquad \frac{\partial \mathcal{L}}{\partial x^\mu} = \frac{1}{2}\partial_\mu g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta.$$

The Euler–Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(g_{\mu\nu}\dot{x}^\nu) - \frac{1}{2}\partial_\mu g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0.$$

Expanding the total derivative and raising with $g^{\rho\mu}$ gives

$$\boxed{\ddot{x}^\rho + \Gamma^\rho{}_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0.}$$

There is no term proportional to $\dot{x}^\rho$, so $\lambda$ is affine. Under a general reparametrization this equation acquires such a term; it retains the same affine form only for $\lambda' = a\lambda + b$. Contracting the geodesic equation with $g_{\rho\sigma}\dot{x}^\sigma$ shows

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta\right) = 0,$$

so the norm of the tangent is constant along the curve and any other affine parameter differs from $\lambda$ only by an affine map $\lambda' = a\lambda + b$. Fixing this constant selects a particular normalization. For a massive particle the choice $\lambda = \tau/m$ gives

$$\dot{x}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}(\tau/m)} = m\,\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = m\,u^\mu = p^\mu,$$

so the tangent *is* the physical four-momentum: this action principle's affine parameter is precisely the one for which $p^\alpha = \mathrm{d}x^\alpha/\mathrm{d}\lambda$, as stated in the problem. Any other affine normalization rescales the tangent to a constant multiple of $p^\mu$.

**Solution to 2 (b).**

The canonical momentum is

$$\boxed{p_\mu = g_{\mu\nu}\dot{x}^\nu.}$$

Inverting gives $\dot{x}^\mu = g^{\mu\nu}p_\nu$. The Legendre transform is

$$\mathcal{H} = p_\mu\dot{x}^\mu - \mathcal{L} = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu - \frac{1}{2}g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \boxed{\frac{1}{2}g^{\mu\nu}p_\mu p_\nu}.$$

The first Hamilton equation gives

$$\dot{x}^\mu = \frac{\partial\mathcal{H}}{\partial p_\mu} = g^{\mu\nu}p_\nu,$$

which reproduces the inverse momentum relation. The second gives

$$\dot{p}_\mu = -\frac{1}{2}(\partial_\mu g^{\alpha\beta})p_\alpha p_\beta = \frac{1}{2}(\partial_\mu g_{\rho\sigma})\dot{x}^\rho\dot{x}^\sigma,$$

where $\partial_\mu g^{\alpha\beta} = -g^{\alpha\rho}g^{\beta\sigma}\partial_\mu g_{\rho\sigma}$ was used. Substituting $p_\mu = g_{\mu\nu}\dot{x}^\nu$ gives

$$g_{\mu\nu}\ddot{x}^\nu + (\partial_\rho g_{\mu\nu})\dot{x}^\rho\dot{x}^\nu - \frac{1}{2}(\partial_\mu g_{\rho\sigma})\dot{x}^\rho\dot{x}^\sigma = 0.$$

Raising the index and symmetrizing the first derivative term produces exactly $\ddot{x}^\mu + \Gamma^\mu{}_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma = 0$. Hence Hamilton's equations and the affine geodesic equation are equivalent. With the same normalization $\lambda = \tau/m$ as in part (a), $p_\mu = mu_\mu$ is the covariant four-momentum; any other affine choice gives the tangent covector up to a constant multiple.

**Solution to 2 (c).**

If the metric is independent of $x^A$, then

$$\frac{\partial \mathcal{H}}{\partial x^A} = \frac{1}{2}(\partial_A g^{\alpha\beta})p_\alpha p_\beta = 0.$$

The second Hamilton equation immediately gives

$$\boxed{\dot{p}_A = -\frac{\partial \mathcal{H}}{\partial x^A} = 0.}$$

Thus a cyclic coordinate produces a conserved covariant momentum. Geometrically, $\partial/\partial x^A$ is a Killing direction and $p_A$ is the associated constant of motion.
Because $\mathcal{H}$ has no explicit $\lambda$ dependence,

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\lambda} = \frac{\partial \mathcal{H}}{\partial x^\mu}\dot{x}^\mu + \frac{\partial \mathcal{H}}{\partial p_\mu}\dot{p}_\mu = 0$$

after inserting Hamilton's equations. Finally, with the normalization of part (a), $p^\mu = mu^\mu$, so for signature $(-,+,+,+)$ $p^\mu p_\mu = -m^2c^2$ and

$$\boxed{\mathcal{H} = -\frac{1}{2}m^2c^2,}$$

constant along the geodesic, as the vanishing of $\mathrm{d}\mathcal{H}/\mathrm{d}\lambda$ already guarantees. For null geodesics the corresponding value is $\mathcal{H} = 0$.

## Exercise 3 — Scalar field in curved spacetime

### Solution to 3 (a).

The Lagrangian density is $\mathcal{L} = -\frac{1}{2}g^{\alpha\beta}\nabla_\alpha\phi\,\nabla_\beta\phi - V(\phi)$. Its derivatives are

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{\mathrm{d}V}{\mathrm{d}\phi}, \qquad \frac{\partial \mathcal{L}}{\partial(\nabla_\mu\phi)} = -\tfrac{1}{2}\big(g^{\mu\beta}\nabla_\beta\phi + g^{\alpha\mu}\nabla_\alpha\phi\big) = -\nabla^\mu\phi,$$

the two terms of the product rule cancelling the factor $\frac{1}{2}$. The Euler–Lagrange equation from the Conventions then gives

$$\frac{\partial \mathcal{L}}{\partial \phi} - \nabla_\mu\frac{\partial \mathcal{L}}{\partial(\nabla_\mu\phi)} = -\frac{\mathrm{d}V}{\mathrm{d}\phi} + \nabla_\mu\nabla^\mu\phi = 0 \quad\Longrightarrow\quad \boxed{\Box\phi - \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0.}$$

*Remark.* Using $\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}V^\mu)$, the wave operator takes the computable form $\Box\phi = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}\,g^{\mu\nu}\partial_\nu\phi)$ —no Christoffel symbols needed. In flat spacetime this reduces to $-\frac{1}{c^2}\ddot{\phi} + \nabla^2\phi = V'$, as it must.

### Solution to 3 (b).

Vary $S_\phi$ with respect to $g^{\alpha\beta}$. The metric appears in two places: explicitly in the kinetic contraction, and in the volume factor $\sqrt{-g}$. Using $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\, g_{\alpha\beta}\delta g^{\alpha\beta}$,

$$\delta S_\phi = \int \Big[-\tfrac{1}{2}\nabla_\alpha\phi\,\nabla_\beta\phi\Big]\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x + \int \Big[-\tfrac{1}{2}g^{\rho\sigma}\nabla_\rho\phi\nabla_\sigma\phi - V\Big]\Big(-\tfrac{1}{2}g_{\alpha\beta}\Big)\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x,$$

so

$$\frac{1}{\sqrt{-g}}\frac{\delta S_\phi}{\delta g^{\alpha\beta}} = -\tfrac{1}{2}\nabla_\alpha\phi\nabla_\beta\phi + \tfrac{1}{4}g_{\alpha\beta}\, g^{\rho\sigma}\nabla_\rho\phi\nabla_\sigma\phi + \tfrac{1}{2}g_{\alpha\beta}V.$$

Multiplying by $-2$,

$$\boxed{T^{(\phi)}_{\alpha\beta} = \nabla_\alpha\phi\,\nabla_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}\, g^{\rho\sigma}\nabla_\rho\phi\nabla_\sigma\phi - g_{\alpha\beta}V(\phi).}$$

*Check.* The general formula of Exercise 5(f), $T_{\alpha\beta} = -2\,\partial\mathcal{L}/\partial g^{\alpha\beta} + g_{\alpha\beta}\mathcal{L}$ with $\partial\mathcal{L}/\partial g^{\alpha\beta} = -\frac{1}{2}\nabla_\alpha\phi\nabla_\beta\phi$, reproduces the same result. Note also $T^{(\phi)}_{00}$ in flat spacetime is $\frac{1}{2}\dot{\phi}^2/c^2 + \frac{1}{2}(\nabla\phi)^2 + V \geq 0$: kinetic plus gradient plus potential energy density, as expected.

**Solution to 3 (c).**

Take the covariant divergence term by term, using $\nabla_\alpha g_{\mu\nu} = 0$ and, for a scalar, $\nabla_\alpha\nabla_\beta\phi = \nabla_\beta\nabla_\alpha\phi$:

$$\nabla^\alpha\big(\nabla_\alpha\phi\,\nabla_\beta\phi\big) = (\Box\phi)\nabla_\beta\phi + \nabla^\alpha\phi\,\nabla_\alpha\nabla_\beta\phi,$$

$$-\tfrac{1}{2}\nabla_\beta\big(\nabla^\sigma\phi\,\nabla_\sigma\phi\big) = -\nabla^\sigma\phi\,\nabla_\beta\nabla_\sigma\phi, \qquad -\nabla_\beta V = -\frac{\mathrm{d}V}{\mathrm{d}\phi}\nabla_\beta\phi.$$

The second term of the first line cancels the second line (relabel $\alpha\to\sigma$ and commute the derivatives), leaving

$$\boxed{\nabla^\alpha T^{(\phi)}_{\alpha\beta} = \Big(\Box\phi - \frac{\mathrm{d}V}{\mathrm{d}\phi}\Big)\nabla_\beta\phi.}$$

The scalar equation of motion therefore implies $\nabla^\alpha T^{(\phi)}_{\alpha\beta} = 0$. Conversely, at any point where $\nabla_\beta\phi \neq 0$, conservation forces the scalar equation of motion. The qualification matters for a constant field configuration, for which the product above can vanish without either factor vanishing separately.

## Exercise 4 — Electromagnetism in curved spacetime

**Solution to 4 (a).**

Write the kinetic term with all metric dependence explicit, $-\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} = -\frac{1}{4}g^{\alpha\mu}g^{\beta\nu}F_{\alpha\beta}F_{\mu\nu}$, where $F_{\alpha\beta}$ (all indices down) is metric-independent, and so is $A_\mu J^\mu$ with the stated index positions. Varying the two inverse metrics gives two equal contributions (relabel dummies and use antisymmetry of $F$):

$$\frac{\partial\mathcal{L}}{\partial g^{\alpha\beta}} = -\tfrac{1}{4}\cdot 2\,F_{\alpha\lambda}F_\beta{}^\lambda = -\tfrac{1}{2}\,F_\alpha{}^\lambda F_{\beta\lambda},$$

which is symmetric in $\alpha\beta$, as it must be. Note that the interaction term $A_\mu J^\mu$ contains no inverse metric (the indices shown are contracted directly), so it does *not* contribute to $\partial\mathcal{L}/\partial g^{\alpha\beta}$; it enters $T_{\alpha\beta}$ only through the volume piece $g_{\alpha\beta}\mathcal{L}$. Then $T_{\alpha\beta} = -2\,\partial\mathcal{L}/\partial g^{\alpha\beta} + g_{\alpha\beta}\mathcal{L}$

(Exercise 5(f)) yields

$$\boxed{T_{\alpha\beta} = F_\alpha{}^\lambda F_{\beta\lambda} - \tfrac{1}{4} g_{\alpha\beta} F^{\rho\sigma} F_{\rho\sigma} + g_{\alpha\beta} A_\lambda J^\lambda.}$$

*Trace.* For $J^\mu = 0$,

$$g^{\alpha\beta} T_{\alpha\beta} = F^{\beta\lambda} F_{\beta\lambda} - \tfrac{1}{4} \cdot 4\, F^{\rho\sigma} F_{\rho\sigma} = 0.$$

Tracelessness reflects the conformal invariance of Maxwell theory in four dimensions—electromagnetism has no intrinsic length scale (the photon is massless). Physically, $T^\mu{}_\mu = 0$ is the statement that radiation has equation of state $p = \rho c^2/3$. As a sanity check, in flat spacetime $T_{00} = \frac{1}{2}(E^2 + B^2)$ in suitable units: the familiar field energy density. (The interaction piece $g_{\alpha\beta} A_\lambda J^\lambda$ is the coupling energy; in a complete treatment it is bundled with the stress–energy of the charged matter that produces $J^\mu$.)

**Solution to 4 (b).**

Vary $A_\nu$ at fixed metric, with $\delta F_{\alpha\beta} = \nabla_\alpha \delta A_\beta - \nabla_\beta \delta A_\alpha$:

$$\delta S = \int \Big[ -\tfrac{1}{2} F^{\alpha\beta} \delta F_{\alpha\beta} + J^\beta \delta A_\beta \Big] \sqrt{-g}\, \mathrm{d}^4 x = \int \Big[ -F^{\alpha\beta} \nabla_\alpha \delta A_\beta + J^\beta \delta A_\beta \Big] \sqrt{-g}\, \mathrm{d}^4 x,$$

using the antisymmetry of $F^{\alpha\beta}$ to combine the two terms. Integrating by parts (the connection is metric compatible, so Stokes applies) and demanding stationarity for arbitrary $\delta A_\beta$,

$$\nabla_\alpha F^{\alpha\beta} + J^\beta = 0 \qquad \Longleftrightarrow \qquad \boxed{\nabla_\mu F^{\nu\mu} = J^\nu.}$$

The homogeneous half of Maxwell's equations, $\nabla_{[\mu} F_{\nu\lambda]} = 0$, holds identically because $F_{\mu\nu}$ is built from a potential.

**Solution to 4 (c).**

Hold $g_{\mu\nu}$ (hence $R^{\mu\nu}$) fixed and vary $A$. Since $\mathcal{L}' = \beta R^{\mu\nu} g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma}$ is symmetric under exchanging the two field strengths (both $R^{\mu\nu}$ and $g^{\rho\sigma}$ are symmetric), its variation picks up a factor of 2:

$$\delta\mathcal{L}' = 2\beta\, R^{\mu\nu} F_\mu{}^\sigma\, \delta F_{\nu\sigma} = 2\beta\, R^{\mu\nu} F_\mu{}^\sigma \big( \nabla_\nu \delta A_\sigma - \nabla_\sigma \delta A_\nu \big).$$

Integrating both terms by parts and relabelling dummy indices, the coefficient of $\delta A_\sigma$ is

$$2\beta \Big[ \nabla_\lambda \big( R^{\mu\sigma} F_\mu{}^\lambda \big) - \nabla_\nu \big( R^{\mu\nu} F_\mu{}^\sigma \big) \Big] = -2\beta\, \nabla_\mu \mathcal{F}^{\sigma\mu}, \qquad \mathcal{F}^{\sigma\mu} \equiv R^{\lambda\mu} F_\lambda{}^\sigma - R^{\lambda\sigma} F_\lambda{}^\mu,$$

with $\mathcal{F}^{\sigma\mu} = -\mathcal{F}^{\mu\sigma}$ antisymmetric by construction. Adding this to the variation of part (b), stationarity now requires $\nabla_\alpha F^{\alpha\nu} + J^\nu - 2\beta \nabla_\mu \mathcal{F}^{\nu\mu} = 0$, i.e.

$$\boxed{\nabla_\mu \big( F^{\nu\mu} + 2\beta\, \mathcal{F}^{\nu\mu} \big) = J^\nu.}$$

The dynamical field strength is shifted by a curvature-dressed piece: in curved regions, spacetime geometry itself polarizes the electromagnetic field. In vacuum GR regions with $R_{\mu\nu} = 0$ the correction disappears, but near matter the propagation of light is modified at order $\beta R$.

**Solution to 4 (d).**

*Einstein's equation.* Write $\mathcal{L}' = \beta R_{\mu\nu}\Theta^{\mu\nu}$ with the symmetric weight $\Theta^{\mu\nu} \equiv F^{\mu\lambda}F^{\nu}{}_{\lambda}$. The metric now appears in *three* places, each generating a distinct kind of term:

- the volume factor, giving $-\frac{1}{2}g_{\alpha\beta}\mathcal{L}'$;
- the explicit inverse metrics hidden in $\Theta^{\mu\nu}$ (three of them, once $F$ is written with lowered indices), giving *algebraic* curvature–field couplings of the schematic form $\beta\, R\, F\, F$ with various contractions;
- the variation $\delta R_{\mu\nu}$ itself. By the Palatini identity this is a total derivative structure, but—exactly as in Exercise 8(a), where the weight was $f(\phi)g^{\mu\nu}$—the non-constant weight $\Theta^{\mu\nu}$ obstructs discarding it. Two integrations by parts land the derivatives on $\Theta^{\mu\nu}$:

$$\int \Theta^{\mu\nu}\delta R_{\mu\nu}\sqrt{-g}\,\mathrm{d}^4x = \int \Big[\tfrac{1}{2}g_{\alpha\beta}\nabla_\mu\nabla_\nu\Theta^{\mu\nu} + \tfrac{1}{2}\Box\Theta_{\alpha\beta} - \nabla_\mu\nabla_{(\alpha}\Theta^{\mu}{}_{\beta)}\Big]\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x,$$

  where $(\alpha\beta)$ denotes symmetrization. (Setting $\Theta^{\mu\nu} = f\,g^{\mu\nu}$ recovers $g_{\alpha\beta}\Box f - \nabla_\alpha\nabla_\beta f$ of Exercise 8(a)—a useful check.)

Einstein's equation therefore becomes $G_{\alpha\beta} = \frac{8\pi G}{c^4}\big(T_{\alpha\beta} + T^{(\beta)}_{\alpha\beta}\big)$, where $T^{(\beta)}_{\alpha\beta}$ contains both algebraic $\beta RFF$ terms and, qualitatively new, *second-derivative* terms $\sim \beta\,\nabla\nabla(FF)$: gravity now responds to gradients of the field-strength bilinears, not just to the local electromagnetic energy density. The equations remain second order in the metric, but the coupling is manifestly non-minimal.
*Current conservation.* The modified Maxwell equation of part (c) has the form $\nabla_\mu H^{\nu\mu} = J^\nu$ with $H^{\nu\mu} = F^{\nu\mu} + 2\beta\mathcal{F}^{\nu\mu}$ *antisymmetric.* Then

$$\nabla_\nu J^\nu = \nabla_\nu\nabla_\mu H^{\nu\mu} = \tfrac{1}{2}[\nabla_\nu, \nabla_\mu]H^{\nu\mu} = \tfrac{1}{2}\big(R_{\lambda\mu}H^{\lambda\mu} + R_{\lambda\nu}H^{\lambda\nu}\big) = R_{\lambda\mu}H^{\lambda\mu} = 0,$$

since the symmetric Ricci tensor is contracted with the antisymmetric $H$. (The first step uses antisymmetry of $H$ to reduce the symmetric part of $\nabla_\nu\nabla_\mu$ to zero; the commutator is evaluated with the Riemann convention of the Conventions.) So

$$\boxed{\nabla_\nu J^\nu = 0\,:}$$

the current is still conserved. This had to happen: the divergence of an antisymmetric tensor is identically divergence-free, and gauge invariance of the $A_\mu J^\mu$ coupling requires a conserved source. The non-minimal term deforms the dynamics but not this consistency structure.

## Exercise 5 — Einstein field equations

### Solution to 5 (a).

Vary $g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha{}_\gamma$: $\delta g^{\alpha\beta}g_{\beta\gamma} + g^{\alpha\beta}\delta g_{\beta\gamma} = 0$, so $g^{\alpha\beta}\delta g_{\beta\gamma} = -\delta g^{\alpha\beta}g_{\beta\gamma}$. Contracting with $g^{\gamma\rho}$ frees the lower indices:

$$\boxed{\delta g_{\alpha\beta} = -g_{\alpha\mu}g_{\beta\nu}\,\delta g^{\mu\nu}.}$$

Contracting gives the trace relation $g_{\alpha\beta}\delta g^{\alpha\beta} = -g^{\alpha\beta}\delta g_{\alpha\beta}$.

### Solution to 5 (b).

Let $\mathbf{M}$ be an invertible square matrix and define $D \equiv \det \mathbf{M}$. Jacobi's formula follows from

$$\ln |D| = \mathrm{Tr}(\ln \mathbf{M}),$$

and gives

$$\frac{\delta D}{D} = \mathrm{Tr}\big(\mathbf{M}^{-1}\delta\mathbf{M}\big)\,.$$

The absolute value is required here because the determinant of a Lorentzian metric is negative. Taking $\mathbf{M} = (g_{\mu\nu})$ and $D = g \equiv \det(g_{\mu\nu})$, the matrix element of $\mathbf{M}^{-1}$ is the inverse metric $(g^{-1})^{\nu\mu} = g^{\nu\mu}$, so $\mathrm{Tr}(\mathbf{M}^{-1}\delta\mathbf{M}) = g^{\nu\mu}\delta g_{\mu\nu}$, and we obtain

$$\frac{\delta g}{g} = g^{\mu\nu}\delta g_{\nu\mu} = g^{\mu\nu}\delta g_{\mu\nu}.$$

Therefore, since $g < 0$,

$$\begin{aligned}
\delta\sqrt{-g} &= -\frac{1}{2\sqrt{-g}}\,\delta g \\
&= -\frac{g}{2\sqrt{-g}}\,g^{\mu\nu}\delta g_{\mu\nu} \\
&= \frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\delta g_{\mu\nu}.
\end{aligned}$$

Hence, if $g_{\mu\nu}$ is taken as the independent variable,

$$\boxed{\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\delta g_{\mu\nu}.}$$

Using the trace relation from part (a), $g^{\mu\nu}\delta g_{\mu\nu} = -g_{\mu\nu}\delta g^{\mu\nu}$, this may also be written as

$$\boxed{\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\,g_{\mu\nu}\delta g^{\mu\nu}.}$$

For completeness, the determinant step can also be verified directly by a cofactor expansion. Let $C^{\mu\nu}$ be the cofactor of the matrix element $g_{\mu\nu}$. To first order,

$$\delta g = C^{\mu\nu}\,\delta g_{\mu\nu}.$$

The inverse-matrix formula gives $(g^{-1})^{\nu\mu} = C^{\mu\nu}/g$. Since the metric is symmetric, $(g^{-1})^{\nu\mu} = g^{\nu\mu} = g^{\mu\nu}$, and therefore

$$C^{\mu\nu} = g\,g^{\mu\nu}, \qquad \delta g = g\,g^{\mu\nu}\delta g_{\mu\nu},$$

which is exactly the determinant identity used above.

**Solution to 5 (c).**

Let the varied connection be $\widetilde{\Gamma}^{\rho}{}_{\mu\nu} = \Gamma^{\rho}{}_{\mu\nu} + \delta\Gamma^{\rho}{}_{\mu\nu}$. Although a connection is not itself a tensor, the difference of two connections is. Hence $\delta\Gamma^{\rho}{}_{\mu\nu}$ is a genuine $(1,2)$ tensor. Since both connections are Levi–Civita connections, it is also symmetric in its lower indices:

$$\delta\Gamma^{\rho}{}_{\mu\nu} = \delta\Gamma^{\rho}{}_{\nu\mu}.$$

With the curvature convention used here, the Ricci tensor is

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

Varying this expression and retaining terms linear in $\delta\Gamma$ gives

$$\begin{aligned}\delta R_{\alpha\beta} = {} & \partial_\mu \delta\Gamma^\mu{}_{\alpha\beta} - \partial_\beta \delta\Gamma^\mu{}_{\alpha\mu} \\ & + \delta\Gamma^\mu{}_{\mu\lambda}\Gamma^\lambda{}_{\alpha\beta} + \Gamma^\mu{}_{\mu\lambda}\delta\Gamma^\lambda{}_{\alpha\beta} \\ & - \delta\Gamma^\mu{}_{\beta\lambda}\Gamma^\lambda{}_{\alpha\mu} - \Gamma^\mu{}_{\beta\lambda}\delta\Gamma^\lambda{}_{\alpha\mu}. \end{aligned} \tag{1}$$

Because $\delta\Gamma^\rho{}_{\mu\nu}$ is a tensor, its covariant derivative is well defined. Expanding the two relevant derivatives gives

$$\begin{aligned}\nabla_\mu \delta\Gamma^\mu{}_{\alpha\beta} &= \partial_\mu \delta\Gamma^\mu{}_{\alpha\beta} + \Gamma^\mu{}_{\mu\lambda}\delta\Gamma^\lambda{}_{\alpha\beta} - \Gamma^\lambda{}_{\mu\alpha}\delta\Gamma^\mu{}_{\lambda\beta} - \Gamma^\lambda{}_{\mu\beta}\delta\Gamma^\mu{}_{\alpha\lambda}, \\ \nabla_\beta \delta\Gamma^\mu{}_{\alpha\mu} &= \partial_\beta \delta\Gamma^\mu{}_{\alpha\mu} + \Gamma^\mu{}_{\beta\lambda}\delta\Gamma^\lambda{}_{\alpha\mu} - \Gamma^\lambda{}_{\beta\alpha}\delta\Gamma^\mu{}_{\lambda\mu} - \Gamma^\lambda{}_{\beta\mu}\delta\Gamma^\mu{}_{\alpha\lambda}.\end{aligned}$$

Subtracting these expressions, the last term in each line cancels because $\Gamma^\lambda{}_{\mu\beta} = \Gamma^\lambda{}_{\beta\mu}$. Using the symmetry of both $\Gamma^\rho{}_{\mu\nu}$ and $\delta\Gamma^\rho{}_{\mu\nu}$ in their lower indices, and relabeling dummy indices, the remaining terms reproduce exactly Eq. (1). Therefore

$$\boxed{\delta R_{\alpha\beta} = \nabla_\mu \delta\Gamma^\mu{}_{\alpha\beta} - \nabla_\beta \delta\Gamma^\mu{}_{\alpha\mu}.}$$

This is the Palatini identity.

**Solution to 5 (d).**

Contract the Palatini identity with $g^{\alpha\beta}$. Metric compatibility, $\nabla_\lambda g^{\alpha\beta} = 0$, allows the inverse metric to be moved inside the covariant derivatives:

$$g^{\alpha\beta}\delta R_{\alpha\beta} = \nabla_\mu\left(g^{\alpha\beta}\delta\Gamma^\mu{}_{\alpha\beta}\right) - \nabla_\beta\left(g^{\alpha\beta}\delta\Gamma^\mu{}_{\alpha\mu}\right).$$

In the second term, relabel the dummy indices according to $\beta \to \mu$, $\mu \to \alpha$, and $\alpha \to \beta$. Using the symmetry of $g^{\alpha\beta}$ and of the lower indices of $\delta\Gamma$, we obtain

$$g^{\alpha\beta}\delta R_{\alpha\beta} = \nabla_\mu\left(g^{\alpha\beta}\delta\Gamma^\mu{}_{\alpha\beta} - g^{\mu\beta}\delta\Gamma^\alpha{}_{\alpha\beta}\right) \equiv \nabla_\mu w^\mu,$$

where

$$w^\mu = g^{\alpha\beta}\delta\Gamma^\mu{}_{\alpha\beta} - g^{\mu\beta}\delta\Gamma^\alpha{}_{\alpha\beta}.$$

For any vector $w^\mu$, $\nabla_\mu w^\mu = (1/\sqrt{-g})\partial_\mu(\sqrt{-g}\, w^\mu)$. Hence

$$\begin{aligned}\int_{\mathcal{M}} g^{\alpha\beta}\delta R_{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x &= \int_{\mathcal{M}} \partial_\mu\left(\sqrt{-g}\, w^\mu\right)\mathrm{d}^4x \\ &= \oint_{\partial\mathcal{M}} w^\mu\,\mathrm{d}\Sigma_\mu,\end{aligned}$$

where $\mathrm{d}\Sigma_\mu = n_\mu\sqrt{|h|}\,\mathrm{d}^3y$ is the directed surface element. Thus the term contributes only at the boundary.

To examine the boundary conditions, use

$$\delta\Gamma^\rho{}_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}\left(\nabla_\mu\delta g_{\sigma\nu} + \nabla_\nu\delta g_{\sigma\mu} - \nabla_\sigma\delta g_{\mu\nu}\right).$$

The conditions $\delta g^{\alpha\beta} = 0$ and $\nabla_\rho\delta g^{\alpha\beta} = 0$ on $\partial\mathcal{M}$ imply the corresponding conditions for $\delta g_{\alpha\beta}$ and hence $\delta\Gamma^\rho{}_{\mu\nu} = 0$ there. Therefore $w^\mu = 0$ on the boundary, and

$$\boxed{\int_{\mathcal{M}} g^{\alpha\beta}\delta R_{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x = 0.}$$

Fixing only the induced metric, $\delta h_{ij} = 0$, does not fix the normal derivatives of the metric and therefore does not force $\delta\Gamma$ to vanish. A well-posed Dirichlet variational principle is obtained by adding the Gibbons–Hawking–York boundary term, whose variation cancels precisely these unwanted normal-derivative terms.

**Solution to 5 (e).**

Vary $S_{\rm EH} = \int g^{\alpha\beta} R_{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x$ in three pieces: $\delta g^{\alpha\beta}\, R_{\alpha\beta}\,\sqrt{-g}$; $g^{\alpha\beta}\delta R_{\alpha\beta}\,\sqrt{-g}$ (a vanishing boundary term, part (d)); and $g^{\alpha\beta} R_{\alpha\beta}\,\delta\sqrt{-g} = -\frac{1}{2} R\, g_{\alpha\beta}\delta g^{\alpha\beta}\,\sqrt{-g}$. Adding the first and third,

$$\delta S_{\rm EH} = \int (R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R)\delta g^{\alpha\beta}\,\sqrt{-g}\,\mathrm{d}^4x \;\Rightarrow\; \boxed{\tfrac{1}{\sqrt{-g}}\,\delta S_{\rm EH}/\delta g^{\alpha\beta} = G_{\alpha\beta}.}$$

**Solution to 5 (f).**

*Cosmological term.* For $S_\Lambda = -2\Lambda \int \sqrt{-g}\,\mathrm{d}^4x$, only $\sqrt{-g}$ varies, so with part (b), $\delta S_\Lambda = -2\Lambda \int (-\frac{1}{2}\sqrt{-g} g_{\alpha\beta}\delta g^{\alpha\beta})\mathrm{d}^4x = \int \Lambda g_{\alpha\beta}\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x$, i.e. $\frac{1}{\sqrt{-g}}\delta S_\Lambda/\delta g^{\alpha\beta} = \Lambda g_{\alpha\beta}$.
*Matter term.* The definition rearranges to $\delta S_{\rm M} = -\frac{1}{2}\int T_{\alpha\beta}\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x$. Varying $S_{\rm M} = \int \mathcal{L}_{\rm M}\sqrt{-g}\,\mathrm{d}^4x$ with $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g} g_{\alpha\beta}\delta g^{\alpha\beta}$,

$$\delta S_{\rm M} = \int \Big(\frac{\partial \mathcal{L}_{\rm M}}{\partial g^{\alpha\beta}} - \tfrac{1}{2}\mathcal{L}_{\rm M} g_{\alpha\beta}\Big)\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x \;\Rightarrow\; \boxed{T_{\alpha\beta} = -2\frac{\partial \mathcal{L}_{\rm M}}{\partial g^{\alpha\beta}} + g_{\alpha\beta}\mathcal{L}_{\rm M}.}$$

**Solution to 5 (g).**

Combining parts (e) and (f),

$$\delta S = \int \big[\kappa(G_{\alpha\beta} + \Lambda g_{\alpha\beta}) - \tfrac{1}{2} T_{\alpha\beta}\big]\delta g^{\alpha\beta}\,\sqrt{-g}\,\mathrm{d}^4x.$$

Requiring $\delta S = 0$ for arbitrary $\delta g^{\alpha\beta}$ forces the bracket to vanish:

$$\boxed{G_{\alpha\beta} + \Lambda g_{\alpha\beta} = \frac{1}{2\kappa} T_{\alpha\beta}.}$$

## Exercise 6 — Fixing $\kappa$

**Solution to 6 (a).**

Contract $G_{\alpha\beta} = \frac{1}{2\kappa} T_{\alpha\beta}$ with $g^{\alpha\beta}$. Since $g^{\alpha\beta} G_{\alpha\beta} = R - 2R = -R$, we get $-R = \frac{1}{2\kappa}T$, i.e. $R = -\frac{1}{2\kappa}T$. Substituting back into $R_{\alpha\beta} = G_{\alpha\beta} + \frac{1}{2} g_{\alpha\beta} R$,

$$\boxed{R_{\alpha\beta} = \frac{1}{2\kappa}\Big(T_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} T\Big).}$$

**Solution to 6 (b).**

For a slow particle the geodesic equation reduces to $\ddot{x}^i \approx -\Gamma^i{}_{00}c^2$. For a static field $\Gamma^i{}_{00} = -\frac{1}{2}g^{ij}\partial_j g_{00} \approx -\frac{1}{2}\partial_i h_{00}$, so $\ddot{x}^i \approx +\frac{c^2}{2}\partial_i h_{00}$. Matching to $\ddot{\boldsymbol{x}} = -\nabla\Phi$ gives $\frac{c^2}{2}h_{00} = -\Phi$, i.e. $h_{00} = -2\Phi/c^2$ and

$$\boxed{g_{00} = -1 + h_{00} \;\Rightarrow\; g_{00} \approx -(1 + 2\Phi/c^2),}$$

the overall sign fixed by requiring an attractive potential well $\Phi < 0$.

**Solution to 6 (c).**

*(i)* To linear order, for a static field the surviving piece of $R_{00}$ is $R_{00} \approx -\frac{1}{2}\nabla^2 h_{00}$. With $h_{00} = -2\Phi/c^2$, $R_{00} \approx +\frac{1}{c^2}\nabla^2\Phi$.
*(ii)* For dust $T_{\alpha\beta} = \rho u_\alpha u_\beta$, $u^\alpha \approx (c, \mathbf{0})$ so $u_0 \approx -c$: $T_{00} \approx \rho c^2$, $T = g^{\alpha\beta}T_{\alpha\beta} = -\rho c^2$. Then $T_{00} - \frac{1}{2}g_{00}T \approx \rho c^2 - \frac{1}{2}(-1)(-\rho c^2) = \frac{1}{2}\rho c^2$.

**Solution to 6 (d).**

The 00 component of part (a), using (c)(i)–(ii), reads

$$\frac{1}{c^2}\nabla^2\Phi = \frac{1}{2\kappa}\cdot\tfrac{1}{2}\rho c^2 \;\Rightarrow\; \nabla^2\Phi = \frac{\rho c^4}{4\kappa}.$$

Comparing with $\nabla^2\Phi = 4\pi G\rho$ gives $c^4/4\kappa = 4\pi G$, hence

$$\boxed{\kappa = \frac{c^4}{16\pi G}.}$$

Then $\frac{1}{2\kappa} = \frac{8\pi G}{c^4}$ and the field equations read $G_{\alpha\beta} + \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta}$.

**Solution to 6 (e).**

*(i)* Move $\Lambda g_{\alpha\beta}$ to the right: $G_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta} - \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4}(T_{\alpha\beta} + T^\Lambda_{\alpha\beta})$ with $T^\Lambda_{\alpha\beta} = -\frac{\Lambda c^4}{8\pi G}g_{\alpha\beta}$. Matching to a perfect fluid $\propto g_{\alpha\beta}$ gives

$$\boxed{\rho_\Lambda c^2 = \frac{\Lambda c^4}{8\pi G}, \qquad p_\Lambda = -\rho_\Lambda c^2} \quad (w = -1).$$

*(ii)* Carrying $\Lambda g_{\alpha\beta}$ through the 00 component (with $g_{00} \approx -1$) adds $-\Lambda c^2$ to the source:

$$\boxed{\nabla^2\Phi = 4\pi G\rho - \Lambda c^2.}$$

For a spherical solution the extra term integrates to $\Phi_\Lambda = -\frac{1}{6}\Lambda c^2 r^2$, giving $\boldsymbol{g}_\Lambda = -\nabla\Phi_\Lambda = +\frac{1}{3}\Lambda c^2\,\boldsymbol{r}$: a repulsion growing linearly with distance for $\Lambda > 0$—the Newtonian face of cosmic acceleration.

## Exercise 7 — Palatini formalism

**Solution to 7 (a).**

Since $\Gamma$ and $\Gamma + \delta\Gamma$ are both connections, $\delta\Gamma$ transforms homogeneously—a $(1,2)$ tensor. Writing $\delta R_{\alpha\beta}$ from the definition and regrouping the ordinary derivatives with the $\Gamma\,\delta\Gamma$ terms into covariant derivatives,

$$\boxed{\delta R_{\alpha\beta} = \nabla_\mu \delta\Gamma^\mu{}_{\alpha\beta} - \nabla_\beta \delta\Gamma^\mu{}_{\alpha\mu},}$$

where $\nabla$ uses the full connection $\Gamma$ (not yet known to be $\tilde{\Gamma}$).

**Solution to 7 (b).**

Let $Q^{\alpha\beta} = \sqrt{-g}\, g^{\alpha\beta}$. Inserting the Palatini identity gives

$$\delta S = \int \mathrm{d}^4x\, Q^{\alpha\beta} \left( \nabla_\lambda \delta\Gamma^\lambda{}_{\alpha\beta} - \nabla_\beta \delta\Gamma^\lambda{}_{\alpha\lambda} \right).$$

Performing the integration by parts with compactly supported variations gives

$$\delta S = \int \mathrm{d}^4x \left[ -(\nabla_\lambda Q^{\alpha\beta}) \delta\Gamma^\lambda{}_{\alpha\beta} + (\nabla_\sigma Q^{\alpha\sigma}) \delta\Gamma^\lambda{}_{\alpha\lambda} \right].$$

Since the torsion-free condition makes the lower indices of $\delta\Gamma^\lambda{}_{\alpha\beta}$ symmetric, the coefficient must also be symmetrized. Thus

$$\delta S = \int \mathrm{d}^4x \left[ -\nabla_\lambda Q^{\alpha\beta} + \delta_\lambda{}^{(\beta} \nabla_\sigma Q^{\alpha)\sigma} \right] \delta\Gamma^\lambda{}_{\alpha\beta},$$

and stationarity gives

$$\boxed{\nabla_\lambda Q^{\alpha\beta} - \delta_\lambda{}^{(\beta} \nabla_\sigma Q^{\alpha)\sigma} = 0.}$$

The covariant derivative here is the derivative appropriate to a tensor density of weight $+1$; this is what makes the integration by parts valid without assuming metric compatibility.

**Solution to 7 (c).**

Define $W^\alpha \equiv \nabla_\sigma Q^{\alpha\sigma}$ and contract the connection equation on $\lambda$ and $\beta$ in $n$ dimensions:

$$W^\alpha - \frac{1}{2} \left( \delta_\beta{}^\beta W^\alpha + \delta_\beta{}^\alpha W^\beta \right) = W^\alpha - \frac{n+1}{2} W^\alpha = -\frac{n-1}{2} W^\alpha = 0.$$

For $n > 1$, $W^\alpha = 0$. Substitution into the uncontracted equation therefore gives

$$\boxed{\nabla_\lambda(\sqrt{-g}\, g^{\alpha\beta}) = 0.}$$

**Solution to 7 (d).**

Expanding the density derivative gives

$$0 = \nabla_\lambda(\sqrt{-g}\, g^{\alpha\beta}) = \sqrt{-g} \left[ \nabla_\lambda g^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} g_{\rho\sigma} \nabla_\lambda g^{\rho\sigma} \right].$$

Contract with $g_{\alpha\beta}$ and define $Q_\lambda = g_{\rho\sigma} \nabla_\lambda g^{\rho\sigma}$. Then

$$\left( 1 - \frac{n}{2} \right) Q_\lambda = 0.$$

Thus $Q_\lambda = 0$ for $n > 2$, and the preceding equation reduces to $\nabla_\lambda g^{\alpha\beta} = 0$. A torsion-free, metric-compatible connection is unique, so

$$\boxed{\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu}(\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}).}$$

Hence the Palatini equation derives the Levi-Civita connection rather than assuming it. This equivalence with the metric formulation need not survive for actions nonlinear in curvature, such as generic Palatini $f(R)$ theories.

## Exercise 8 — Scalar–tensor gravity

### Solution to 8 (a).

The variation proceeds as in Einstein–Hilbert, but $f(\phi)$ rides along as a fixed weight (independent of $g^{\alpha\beta}$). The $\delta R$ piece produces, besides $f R_{\alpha\beta}\,\delta g^{\alpha\beta}$, the derivative terms $\int f[g_{\alpha\beta}\Box - \nabla_\alpha\nabla_\beta]\delta g^{\alpha\beta}\sqrt{-g}\,\mathrm{d}^4x$. For constant $f$ these are total derivatives; for non-constant $f$ we integrate by parts twice, landing the derivatives on $f$: $\int f g_{\alpha\beta}\Box\delta g^{\alpha\beta} \to \int(\Box f)g_{\alpha\beta}\delta g^{\alpha\beta}$ and $-\int f\nabla_\alpha\nabla_\beta\delta g^{\alpha\beta} \to -\int(\nabla_\alpha\nabla_\beta f)\delta g^{\alpha\beta}$. Adding $fG_{\alpha\beta}$,

$$\boxed{\delta S_g = \int \sqrt{-g}\,\mathrm{d}^4x\,[fG_{\alpha\beta} + g_{\alpha\beta}\Box f - \nabla_\alpha\nabla_\beta f]\delta g^{\alpha\beta}.}$$

It is the double integration by parts that generates $\Box f$ and $\nabla_\alpha\nabla_\beta f$; they are absent in GR only because there $f$ is constant.

### Solution to 8 (b).

Setting $\delta(S_g + S_\phi + S_\mathrm{M}) = 0$ and using $T^{(i)}_{\alpha\beta} = -\frac{2}{\sqrt{-g}}\delta S_i/\delta g^{\alpha\beta}$,

$$\boxed{G_{\alpha\beta} = f^{-1}\left[\tfrac{1}{2}T^\mathrm{M}_{\alpha\beta} + \tfrac{1}{2}T^\phi_{\alpha\beta} + \nabla_\alpha\nabla_\beta f - g_{\alpha\beta}\Box f\right],}$$

with (varying $S_\phi$) $T^\phi_{\alpha\beta} = h\nabla_\alpha\phi\nabla_\beta\phi - g_{\alpha\beta}[\frac{1}{2}hg^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi + V]$. **Interpretation.** The coefficient of $T^\mathrm{M}_{\alpha\beta}$ is $\frac{1}{2}f^{-1}$, playing the role of $8\pi G/c^4$. Comparing with GR fixes $f = c^4/(16\pi G)$ for constant $\phi$; when $\phi$ varies slowly, the effective coupling is $G_\mathrm{eff} = c^4/[16\pi f(\phi)]$.

### Solution to 8 (c).

Varying with respect to $\phi$: the kinetic term gives (after parts) $h\Box\phi + \frac{1}{2}h' g^{\alpha\beta}\nabla_\alpha\phi\nabla_\beta\phi$, the potential gives $-V'$, and $S_g$ contributes $Rf'$ since $f = f(\phi)$ multiplies $R$. The matter action is $\phi$-independent. Summing to zero,

$$\boxed{h\Box\phi + \tfrac{1}{2}\frac{\mathrm{d}h}{\mathrm{d}\phi}g^{\alpha\beta}\nabla_\alpha\phi\nabla_\beta\phi - \frac{\mathrm{d}V}{\mathrm{d}\phi} + R\frac{\mathrm{d}f}{\mathrm{d}\phi} = 0.}$$

**Comment.** The $Rf'$ term is the signature of scalar–tensor gravity: curvature sources the scalar, so $\phi$ is not a passive field on a fixed geometry—the geometry drives it and (through

$f$) it back-reacts on the strength of gravity. A minimally coupled scalar ($f$ =const) loses this term.

**Solution to 8 (d).**

*(i)* Substituting $f = c^4\phi/16\pi$, $h = c^4\omega/8\pi\phi$, and $V = 0$:

$$S_{\rm BD} = \frac{c^4}{16\pi}\int \mathrm{d}^4x\,\sqrt{-g}\left[\phi R - \frac{\omega}{\phi}g^{\alpha\beta}\nabla_\alpha\phi\nabla_\beta\phi\right],$$

plus $S_{\rm M}$.
*(ii)* From $f = c^4/(16\pi G_{\rm eff}) = c^4\phi/(16\pi)$,

$$\boxed{G_{\rm eff}(\phi) = \frac{1}{\phi}.}$$

The scalar field literally is the inverse gravitational coupling.
*(iii)* Combining the $\phi$ equation with the trace of the gravitational equation to eliminate $R$ gives the standard Brans–Dicke wave equation $\Box\phi = \frac{8\pi}{c^4(3+2\omega)}T^{\rm M}$. As $\omega \to \infty$ the source $\propto 1/\omega \to 0$, so $\Box\phi \to 0$ and (with suitable boundary conditions) $\phi \to$const, $\nabla\phi \to 0$. The scalar freezes, $G_{\rm eff} \to$const, and the equations reduce to $G_{\alpha\beta} = (8\pi G/c^4)T^{\rm M}_{\alpha\beta}$: ordinary GR. Thus $1/\omega$ measures the departure from GR, and solar-system bounds ($\omega \gtrsim 40{,}000$) keep Brans–Dicke close to Einstein gravity.

## Exercise 9 — Kaluza–Klein compactification

**Solution to 9 (a).**

In the coordinates $(x^\alpha, y^i)$ the metric (5) is block diagonal, with blocks $g_{\alpha\beta}$ and $b^2\gamma_{ij}$. The determinant of a block-diagonal matrix is the product of the block determinants, and $\det(b^2\gamma_{ij}) = b^{2d}\det\gamma_{ij}$ for a $d\times d$ block. Hence

$$\det G_{ab} = \det g_{\alpha\beta}\cdot b^{2d}\det\gamma_{ij} = b^{2d}\,g\,\gamma,$$

and since $g < 0$, $\gamma > 0$, $b > 0$,

$$\boxed{\sqrt{-G} = b^d\sqrt{\gamma}\,\sqrt{-g}.}$$

**Solution to 9 (b).**

Apply the Levi-Civita formula $\Gamma^c{}_{ab} = \frac{1}{2}G^{cd}(\partial_a G_{db} + \partial_b G_{da} - \partial_d G_{ab})$ to the block-diagonal metric with $G_{\alpha\beta} = g_{\alpha\beta}(x)$, $G_{ij} = b^2(x)\gamma_{ij}(y)$, $G^{\alpha\beta} = g^{\alpha\beta}$, $G^{ij} = b^{-2}\gamma^{ij}$, and remember $\partial_i b = 0$, $\partial_\alpha\gamma_{ij} = 0$, $\partial_i g_{\alpha\beta} = 0$.
*Pure blocks.* For all indices four-dimensional only $g_{\alpha\beta}$ enters, so $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\beta\gamma}[g]$. For all indices internal, the factors $b^2$ in $G^{il}$ and $\partial G_{lj}$ cancel (as $\partial_k b = 0$), leaving $\Gamma^i{}_{jk} = \Gamma^i{}_{jk}[\gamma]$.
*Mixed components.*

$$\Gamma^\alpha{}_{ij} = \tfrac{1}{2}g^{\alpha\nu}\big(\partial_i G_{\nu j} + \partial_j G_{\nu i} - \partial_\nu G_{ij}\big) = -\tfrac{1}{2}g^{\alpha\nu}\partial_\nu(b^2)\,\gamma_{ij} = -\,b\,(\nabla^\alpha b)\,\gamma_{ij},$$

$$\Gamma^i{}_{\alpha j} = \tfrac{1}{2}b^{-2}\gamma^{ik}\big(\partial_\alpha G_{kj} + \partial_j G_{k\alpha} - \partial_k G_{\alpha j}\big) = \tfrac{1}{2}b^{-2}\gamma^{ik}\partial_\alpha(b^2)\,\gamma_{kj} = b^{-1}(\nabla_\alpha b)\,\delta^i{}_j.$$

The remaining mixed components vanish because every term involves either an off-block metric component ($G_{\alpha i} = 0$), a $y$-derivative of $g_{\alpha\beta}$, or an $x$-derivative acting where only $\gamma_{ij}(y)$ appears with all-internal free indices: $\Gamma^i{}_{\alpha\beta} = -\frac{1}{2} b^{-2}\gamma^{ik}\partial_k G_{\alpha\beta} = 0$ and $\Gamma^\alpha{}_{\beta i} = \frac{1}{2} g^{\alpha\nu}\partial_i G_{\nu\beta} = 0$.

**Solution to 9 (c).**

With our conventions, $R_{ab} = R^c{}_{acb} = \partial_c \Gamma^c{}_{ba} - \partial_b \Gamma^c{}_{ca} + \Gamma^c{}_{ce}\Gamma^e{}_{ba} - \Gamma^c{}_{be}\Gamma^e{}_{ca}$.
*Four-dimensional block.* Setting $(a, b) = (\alpha, \beta)$, the terms built purely from $\Gamma[g]$ assemble into $R[g]_{\alpha\beta}$. The extra contributions involve $\Gamma^i{}_{\alpha j} = b^{-1}(\nabla_\alpha b)\delta^i{}_j$, whose trace is $\Gamma^i{}_{i\alpha} = d\, b^{-1}\nabla_\alpha b$:

$$-\partial_\beta \Gamma^i{}_{i\alpha} + \Gamma^i{}_{i\lambda}\Gamma^\lambda{}_{\beta\alpha} - \Gamma^i{}_{\beta j}\Gamma^j{}_{i\alpha} = -d\,\partial_\beta\big(b^{-1}\partial_\alpha b\big) + d\, b^{-1}(\partial_\lambda b)\Gamma^\lambda{}_{\beta\alpha} - d\, b^{-2}(\partial_\alpha b)(\partial_\beta b).$$

Expanding $\partial_\beta(b^{-1}\partial_\alpha b) = b^{-1}\partial_\alpha\partial_\beta b - b^{-2}(\partial_\alpha b)(\partial_\beta b)$, the $b^{-2}\partial b\,\partial b$ pieces cancel and the rest combines into a covariant second derivative:

$$\boxed{R_{\alpha\beta} = R[g]_{\alpha\beta} - d\, b^{-1}\nabla_\alpha \nabla_\beta b.}$$

*Internal block.* Setting $(a, b) = (i, j)$, the terms built purely from $\Gamma[\gamma]$ assemble into $R[\gamma]_{ij}$. The extra contributions are

$$\partial_\mu \Gamma^\mu{}_{ij} + \Gamma^\mu{}_{\mu\lambda}\Gamma^\lambda{}_{ij} + \Gamma^k{}_{k\lambda}\Gamma^\lambda{}_{ij} - \Gamma^\mu{}_{ik}\Gamma^k{}_{\mu j} - \Gamma^k{}_{i\mu}\Gamma^\mu{}_{kj}.$$

Substituting the mixed Christoffels:

$$\partial_\mu \Gamma^\mu{}_{ij} + \Gamma^\mu{}_{\mu\lambda}\Gamma^\lambda{}_{ij} = -\gamma_{ij}\big[\partial_\mu(b\,\nabla^\mu b) + \Gamma^\mu{}_{\mu\lambda}\, b\,\nabla^\lambda b\big] = -\gamma_{ij}\,\nabla_\mu(b\,\nabla^\mu b) = -\gamma_{ij}\big[b\,\Box b + (\nabla b)^2\big],$$

$$\Gamma^k{}_{k\lambda}\Gamma^\lambda{}_{ij} = d\, b^{-1}(\nabla_\lambda b)\big(-b\,\nabla^\lambda b\,\gamma_{ij}\big) = -d\,(\nabla b)^2\gamma_{ij},$$

$$-\Gamma^\mu{}_{ik}\Gamma^k{}_{\mu j} - \Gamma^k{}_{i\mu}\Gamma^\mu{}_{kj} = +(\nabla b)^2\gamma_{ij} + (\nabla b)^2\gamma_{ij} = +2(\nabla b)^2\gamma_{ij}.$$

Adding everything,

$$\boxed{R_{ij} = R[\gamma]_{ij} - \gamma_{ij}\big[b\,\Box b + (d-1)(\nabla b)^2\big].}$$

**Solution to 9 (d).**

The inverse metric is $G^{\alpha\beta} = g^{\alpha\beta}$, $G^{ij} = b^{-2}\gamma^{ij}$, so $R[G] = g^{\alpha\beta}R_{\alpha\beta} + b^{-2}\gamma^{ij}R_{ij}$. Using $\gamma^{ij}\gamma_{ij} = d$ and part (c),

$$R[G] = R[g] - d\, b^{-1}\Box b + b^{-2}R[\gamma] - d\, b^{-2}\big[b\,\Box b + (d-1)(\nabla b)^2\big],$$

i.e.

$$\boxed{R[G] = R[g] + b^{-2}R[\gamma] - 2d\, b^{-1}\Box b - d(d-1)\, b^{-2}(\nabla b)^2.}$$

**Solution to 9 (e).**

Insert $\sqrt{-G} = b^d\sqrt{\gamma}\sqrt{-g}$ (part (a)) and the curvature decomposition (part (d)) into (6). Nothing in the integrand except $\sqrt{\gamma}$ depends on $y^i$, so $\int\sqrt{\gamma}\,\mathrm{d}^d y = \mathcal{V}$ factors out; using

$\mathcal{V}/G_{4+d} = 1/G_4$ and $R[\gamma] = d(d-1)k$,

$$S = \int \mathrm{d}^4x \sqrt{-g}\left\{\frac{c^4}{16\pi G_4}\Big[b^d R[g] + d(d-1)k\, b^{d-2} - 2d\, b^{d-1}\Box b - d(d-1)b^{d-2}(\nabla b)^2\Big] + \mathcal{V} b^d \mathcal{L}_{\mathrm{M}}\right\}.$$

Integrate the $\Box b$ term by parts:

$$-2d\int \sqrt{-g}\, b^{d-1}\Box b = +2d\int \sqrt{-g}\,(\nabla_\mu b^{d-1})\nabla^\mu b = 2d(d-1)\int \sqrt{-g}\, b^{d-2}(\nabla b)^2,$$

dropping the total derivative $\nabla_\mu(b^{d-1}\nabla^\mu b)$. The net kinetic coefficient is $2d(d-1) - d(d-1) = +d(d-1)$:

$$S = \int \mathrm{d}^4x \sqrt{-g}\Big\{\frac{c^4}{16\pi G_4}\Big[b^d R[g] + d(d-1)b^{d-2}g^{\alpha\beta}(\nabla_\alpha b)(\nabla_\beta b) + d(d-1)k\, b^{d-2}\Big] + \mathcal{V} b^d \mathcal{L}_{\mathrm{M}}\Big\}.$$

**Solution to 9 (f).**

With $\Phi = b^d$ we have $b^d R[g] = \Phi R[g]$, and $\nabla_\alpha \Phi = d\, b^{d-1}\nabla_\alpha b$ gives $(\nabla\Phi)^2 = d^2 b^{2d-2}(\nabla b)^2$, i.e. $b^{d-2}(\nabla b)^2 = (\nabla\Phi)^2/(d^2\Phi)$. Hence

$$d(d-1)\, b^{d-2}(\nabla b)^2 = \frac{d-1}{d}\,\frac{(\nabla\Phi)^2}{\Phi} = -\,\omega_{\mathrm{BD}}\,\frac{(\nabla\Phi)^2}{\Phi}, \qquad \boxed{\omega_{\mathrm{BD}} = -\frac{d-1}{d}.}$$

Thus (up to the constant normalization of the scalar and the curvature term, which acts as a potential) Kaluza–Klein reduction produces a Brans–Dicke theory with $\omega_{\mathrm{BD}} \in (-1, 0]$: $\omega_{\mathrm{BD}} = 0$ for $d = 1$ and $\omega_{\mathrm{BD}} \to -1$ as $d \to \infty$. This is an $\mathcal{O}(1)$ value, in catastrophic conflict with the solar-system bound $\omega \gtrsim 4 \times 10^4$ of Exercise 8(d)—so a Kaluza–Klein radion must *not* remain a massless Brans–Dicke-type scalar; it must acquire a mass (be stabilized) to suppress the long-range fifth force.

**Solution to 9 (g).**

Write $b = e^\beta$; then $b^{d-2}(\nabla b)^2 = e^{d\beta}(\nabla\beta)^2$ and the Jordan-frame action of part (e) reads

$$S = \int \mathrm{d}^4x \sqrt{-g}\Big\{\frac{c^4}{16\pi G_4}\Big[e^{d\beta}R[g] + d(d-1)e^{d\beta}(\nabla\beta)^2 + d(d-1)k\, e^{(d-2)\beta}\Big] + \mathcal{V} e^{d\beta}\mathcal{L}_{\mathrm{M}}\Big\}.$$

Set $\tilde{g}_{\mu\nu} = e^{2\omega}g_{\mu\nu}$ with $2\omega = d\beta$, so that

$$\sqrt{-g} = e^{-2d\beta}\sqrt{-\tilde{g}}, \qquad g^{\mu\nu} = e^{d\beta}\,\tilde{g}^{\mu\nu}, \qquad R[g] = e^{d\beta}\big(R[\tilde{g}] + 3d\,\tilde{\Box}\beta - \tfrac{3d^2}{2}(\tilde{\nabla}\beta)^2\big),$$

the last from the given identity with $\omega = d\beta/2$. Term by term:

- Curvature: $\sqrt{-g}\, e^{d\beta}R[g] = e^{-2d\beta}e^{d\beta}e^{d\beta}\sqrt{-\tilde{g}}\,(R[\tilde{g}] + 3d\,\tilde{\Box}\beta - \tfrac{3d^2}{2}(\tilde{\nabla}\beta)^2) = \sqrt{-\tilde{g}}\,(R[\tilde{g}] - \tfrac{3d^2}{2}(\tilde{\nabla}\beta)^2)$, after dropping the total derivative $\sqrt{-\tilde{g}}\,\tilde{\Box}\beta = \partial_\mu(\sqrt{-\tilde{g}}\,\tilde{\nabla}^\mu\beta)$.
- Kinetic: $(\nabla\beta)^2_g = g^{\mu\nu}\partial_\mu\beta\,\partial_\nu\beta = e^{d\beta}(\tilde{\nabla}\beta)^2$, so $\sqrt{-g}\, d(d-1)e^{d\beta}(\nabla\beta)^2_g = \sqrt{-\tilde{g}}\, d(d-1)(\tilde{\nabla}\beta)^2$.
- Internal curvature: $\sqrt{-g}\, d(d-1)k\, e^{(d-2)\beta} = \sqrt{-\tilde{g}}\, d(d-1)k\, e^{-(d+2)\beta}$.

- Matter: after writing the original metric as $g_{\mu\nu} = e^{-d\beta}\tilde{g}_{\mu\nu}$,

$$\sqrt{-g}\, e^{d\beta}\mathcal{V}\mathcal{L}_\mathrm{M}(g,\psi) = \sqrt{-\tilde{g}}\,\mathcal{V}e^{-d\beta}\mathcal{L}_\mathrm{M}(e^{-d\beta}\tilde{g},\psi).$$

The kinetic terms combine as $-\frac{3d^2}{2} + d(d-1) = -\frac{1}{2}d(d+2)$, giving

$$S = \int \mathrm{d}^4x\,\sqrt{-\tilde{g}}\Big\{\frac{c^4}{16\pi G_4}\Big[R[\tilde{g}] - \tfrac{1}{2}d(d+2)(\tilde{\nabla}\beta)^2 + d(d-1)k\,e^{-(d+2)\beta}\Big] \\ + \mathcal{V}e^{-d\beta}\mathcal{L}_\mathrm{M}(e^{-d\beta}\tilde{g},\psi)\Big\}.$$

Note the *negative* exponent $e^{-(d+2)\beta}$: the internal curvature term is redshifted, not enhanced, by growing extra dimensions.

**Solution to 9 (h).**

Since $\frac{c^4}{16\pi G_4} = \frac{1}{2}\bar{m}_\mathrm{P}^2$, the $\beta$-kinetic term is

$$-\frac{c^4}{16\pi G_4}\cdot\tfrac{1}{2}d(d+2)(\tilde{\nabla}\beta)^2 = -\tfrac{1}{4}d(d+2)\,\bar{m}_\mathrm{P}^2(\tilde{\nabla}\beta)^2 = -\tfrac{1}{2}(\tilde{\nabla}\phi)^2 \quad \text{for} \quad \phi = \sqrt{\tfrac{d(d+2)}{2}}\,\bar{m}_\mathrm{P}\,\beta,$$

so $\phi$ is canonically normalized. Substituting $\beta = \phi/\big(\sqrt{d(d+2)/2}\,\bar{m}_\mathrm{P}\big)$ into the exponentials,

$$(d+2)\beta = \sqrt{\frac{2(d+2)}{d}}\,\frac{\phi}{\bar{m}_\mathrm{P}}, \qquad d\beta = \sqrt{\frac{2d}{d+2}}\,\frac{\phi}{\bar{m}_\mathrm{P}},$$

and $\frac{c^4}{16\pi G_4}\,d(d-1)k = \frac{1}{2}d(d-1)k\,\bar{m}_\mathrm{P}^2$, so

$$S = \int \mathrm{d}^4x\sqrt{-\tilde{g}}\Big\{\frac{c^4}{16\pi G_4}R[\tilde{g}] - \tfrac{1}{2}(\tilde{\nabla}\phi)^2 + \tfrac{1}{2}d(d-1)k\,\bar{m}_\mathrm{P}^2 e^{-\sqrt{2(d+2)/d}\,\phi/\bar{m}_\mathrm{P}} \\ + \mathcal{V}e^{-\sqrt{2d/(d+2)}\,\phi/\bar{m}_\mathrm{P}}\mathcal{L}_\mathrm{M}(e^{-d\beta}\tilde{g},\psi)\Big\}.$$

Ignoring matter, the radion potential (the negative of the non-kinetic scalar term) is

$$V(\phi) = -\tfrac{1}{2}\,d(d-1)\,k\,\bar{m}_\mathrm{P}^2\,e^{-\sqrt{2(d+2)/d}\,\phi/\bar{m}_\mathrm{P}}.$$

Note also the exponential coupling of $\phi$ to $\mathcal{L}_\mathrm{M}$: in the Einstein frame matter feels the radion directly, with gravitational strength.

**Solution to 9 (i).**

Write $V(\phi) = -\frac{1}{2}d(d-1)k\,\bar{m}_\mathrm{P}^2\,e^{-\lambda\phi/\bar{m}_\mathrm{P}}$ with $\lambda = \sqrt{2(d+2)/d} > 0$, and recall $\phi \propto \beta = \log b$. *Flat internal space, $k = 0$:* the potential vanishes identically and the radion is exactly massless. It then mediates a long-range, gravitational-strength fifth force via its coupling to $\mathcal{L}_\mathrm{M}$—equivalently, the theory is Brans–Dicke with $\omega_\mathrm{BD} = -(d-1)/d$ (part (f))—which is excluded by solar-system tests.
*Positive curvature, $k > 0$ (e.g. a sphere):* $V < 0$ and $V'(\phi) = +\frac{1}{2}d(d-1)k\lambda\,\bar{m}_\mathrm{P}\,e^{-\lambda\phi/\bar{m}_\mathrm{P}} > 0$, so $V$ increases monotonically toward $0^-$; the field rolls downhill to $\phi \to -\infty$, where $V \to -\infty$. Since $\phi \propto \log b$, this is $b \to 0$: the extra dimensions collapse.

*Negative curvature, $k < 0$:* $V > 0$ and monotonically decreasing toward $0^+$; the field rolls to $\phi \to +\infty$, i.e. $b \to \infty$: the extra dimensions decompactify.
In no case does $V$ possess a minimum, so the ansatz by itself has no stable compactified size.
**Conclusion:** stability requires additional contributions to the potential—an appropriate matter Lagrangian and field configuration in the extra dimensions (in modern language, fluxes or other stabilization mechanisms)—to generate a minimum in $V(\phi)$ and give the radion a mass, which simultaneously evades the fifth-force problem of part (f).