Kinetic Energy as Potential

The Idea: A force is *conservative* if it is the (negative) gradient of a scalar-valued function. For example, the gravitational force acting on a particle of mass m near the surface of the earth is

$$\mathsf{f}_g = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} = - \begin{bmatrix} \frac{\partial mgz}{\partial x} \\ \frac{\partial mgz}{\partial y} \\ \frac{\partial mgz}{\partial z} \end{bmatrix} = -\nabla \underbrace{mgz}_{\Phi}.$$

The scalar-valued function $\Phi(x, y, z) = mgz$ is called a *potential* (energy) for f_g. The gravitational potential $\Phi(x, y, z) = mgz$ is a multiple of the altitude z, so the higher the particle, the greater its potential energy.

Gradients point in the direction of greatest increase of the potential — "up", in the case of the gravitational potential. The minus sign in $\mathsf{f}_g = -\nabla\Phi$ says the force pulls in the opposite direction — "down", in the case of the gravitational potential.

Not all forces are conservative. Non-conservative forces include forces of friction and external stimuli like parents pushing a swing. Any force that dissipates or supplies energy is likely to be non-conservative.

Newton's second law,

$$\mathsf{f} = \frac{d}{dt}(m\mathsf{v}), \tag{1}$$

says that $\frac{d}{dt}(m\mathsf{v})$ is the net force f acting on the particle. Lagrange asks in what sense $\frac{d}{dt}(m\mathsf{v})$ is conservative. His answer is the Lagrangian formulation of mechanics. The potential for $\frac{d}{dt}(m\mathsf{v})$ is the kinetic energy.

These notes derive (non-variational) Lagrangian mechanics from Newtonian mechanics. The derivation clarifies how the sometimes-confusing gradients are interpreted. The proofs require only the simplest tools of analysis: the linearity of the derivative, the product and chain rules, some elementary linear algebra, and the equality of mixed partial derivatives.

Newtonian framework: Newtonian mechanics is governed by Newton's second law, Equation (1). The net force f acting on a particle of mass m changes the momentum $\mathsf{p} = m\mathsf{v}$. The velocity $\mathsf{v} = \dot{\mathsf{x}}$ is the time derivative of position x — the "dot" means "derivative with respect to time". When m is constant, the second law is "f equals $m\mathsf{a}$", but the general law is "f equals $\dot{\mathsf{p}}$".

Suppose, in an effort to solve a particular problem, we make a change of variables

$$\mathsf{x} = \mathsf{G}(\mathsf{q}).$$

Physicists call q the *generalized coordinates* and x the *system* vector.

Example: A spherical pendulum is a point mass m, called the *bob*, suspended by a string (or rigid but massless rod) of fixed length r. The string restricts the bob's motion to a sphere, suggesting a change of variables to spherical coordinates

$$\mathsf{x} = \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix} = \mathsf{G}\left(\begin{bmatrix} \theta \\ \varphi \end{bmatrix}\right).$$

The parameter $\mathsf{q} = \begin{bmatrix} \theta \\ \varphi \end{bmatrix}$ is 2-dimensional; the system vector x is 3-dimensional.

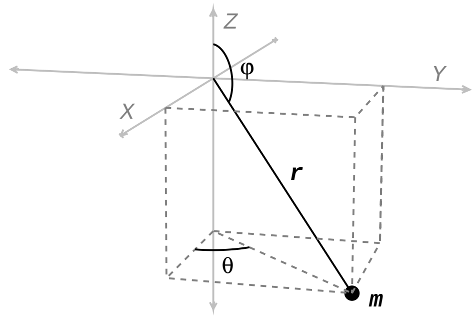

FIGURE 1: SPHERICAL PENDULUM

The chain rule says that the bob's velocity is

$$\mathsf{v} = \dot{\mathsf{x}} = \begin{bmatrix} -r\sin(\theta)\dot{\theta}\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot{\varphi} \\ r\cos(\theta)\dot{\theta}\sin(\varphi) + r\sin(\theta)\cos(\varphi)\dot{\varphi} \\ -r\sin(\varphi)\dot{\varphi} \end{bmatrix} = \underbrace{\begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}}_{\mathsf{S}(\mathsf{q})} \underbrace{\begin{bmatrix} \dot{\theta} \\ \dot{\varphi} \end{bmatrix}}_{\dot{\mathsf{q}}}. \tag{2}$$

The matrix $\mathsf{S} = \mathsf{S}(\mathsf{q})$ is the Jacobian matrix of partial derivatives of G:

$$\mathsf{S}(\mathsf{q}) = \mathbf{D}_\mathsf{q}\mathsf{G}(\mathsf{q}) = \begin{bmatrix} \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}.$$

Under the change of variables, $\mathsf{v} = \mathsf{S}\dot{\mathsf{q}}$ and Newton's second law, Equation (1), becomes

$$\mathsf{f} = \frac{d}{dt}\left(\mathsf{MS}\dot{\mathsf{q}}\right), \tag{3}$$

where M is the *mass matrix*. The spherical pendulum, with its single point mass, has $\mathsf{M} = m\mathsf{I}_3$, the scalar m times the 3×3 identity matrix. M need not be diagonal, but, in Newtonian mechanics, it is symmetric.

The key[1] to the Newtonian framework is to left-multiply both sides of Equation (3) by S^T:

$$\mathsf{S}^T\mathsf{f} = \mathsf{S}^T\frac{d}{dt}\left(\mathsf{MS}\dot{\mathsf{q}}\right). \tag{4}$$

The point is to eliminate the forces of constraint on the left side. The argument goes like this:

1. $\mathsf{S} = \mathsf{S}(\mathsf{q})$ maps *generalized* velocities $\dot{\mathsf{q}}$ to *system* velocities $\mathsf{v} = \dot{\mathsf{x}} = \mathsf{S}\dot{\mathsf{q}}$, so
2. The range (or column space) of $\mathsf{S}(\mathsf{q})$ lies in the tangent plane to the constraint surface at $\mathsf{x} = \mathsf{G}(\mathsf{q})$, so
3. The columns of $\mathsf{S}(\mathsf{q})$ are orthogonal to any vector normal to the constraint surface at $\mathsf{x} = \mathsf{G}(\mathsf{q})$.

In matrix language, $\mathsf{S}^T\mathsf{n} = \mathsf{0}$ for any vector n normal to the constraint surface at $\mathsf{x} = \mathsf{G}(\mathsf{q})$.

In Newtonian mechanics, constraining forces are normal to the constraint surface. For example, the spherical pendulum's string pulls "up" in a direction parallel to x — normal to the sphere. Since

$$\mathsf{S}^T\mathsf{x} = \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\sin(\varphi) & 0 \\ r\cos(\theta)\cos(\varphi) & r\sin(\theta)\cos(\varphi) & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and since the force of constraint f_c is parallel to x,

$$\mathsf{S}^T\mathsf{f}_c = \mathsf{0} \quad \text{for the force of constraint } \mathsf{f}_c.$$

This assumption is called the *principle of virtual work*, or *Newton's third law*, or[2] *Postulate A*.

Finally, forces add vectorially[3] in Newtonian mechanics. We decompose the net force acting on the particle into "applied" forces f_a, such as gravity, and constraint forces f_c, such as the tension in the string:

$$\mathsf{f} = \mathsf{f}_a + \mathsf{f}_c \qquad \text{so} \qquad \mathsf{S}^T\mathsf{f} = \mathsf{S}^T\mathsf{f}_a + \mathsf{S}^T\mathsf{f}_c = \mathsf{S}^T\mathsf{f}_a + \mathsf{0}.$$

Only the applied force remains — the force of constraint is eliminated. Equation (4) becomes

$$\mathsf{S}^T\mathsf{f}_a = \mathsf{S}^T\frac{d}{dt}\left(\mathsf{MS}\dot{\mathsf{q}}\right), \tag{5}$$

with (the known) f_a replacing (the unknown) f on the left side.

It would be nice to see "the derivative of something" rather than "S^T times the derivative of something" on the right side of Equation (5). The product rule is just the ticket:

$$\mathsf{S}^T\mathsf{f}_a = \frac{d}{dt}\left(\mathsf{S}^T\mathsf{MS}\dot{\mathsf{q}}\right) - \dot{\mathsf{S}}^T\mathsf{MS}\dot{\mathsf{q}}. \tag{6}$$

[1] see http://joelshapiro.org/Pubvit/Downloads/Rulla_Newton_VarPrinc.pdf
[2] Cornelius Lanczos, *The Variational Principles of Mechanics*, 4th ed., Dover, 1970, p. 76.
[3] See the notes in footnote 1 for a discussion.

Equation (6) is the *Newtonian framework*. The differentiated variable $\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}$ is the *generalized momentum*, and is confusingly denoted p, the same notation as the (non-generalized) momentum. The components of $\mathsf{S}^T\mathsf{f}_a$ are called *generalized forces*, even when they are not measured in units of force.

Example (spherical pendulum, continued): The forces acting on the bob are dominated by the (applied) force of gravity pulling down, and the (constraint) force of the string pulling "up", toward the pivot. The downward force of gravity is easily modeled by

$$\mathsf{f}_a = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix},$$

where m is the mass of the bob and $g \approx 9.8\frac{\text{m}}{\text{s}^2}$. The generalized force is therefore

$$\mathsf{S}^T\mathsf{f}_a = \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\sin(\varphi) & 0 \\ r\cos(\theta)\cos(\varphi) & r\sin(\theta)\cos(\varphi) & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} = \begin{bmatrix} 0 \\ mgr\sin(\varphi) \end{bmatrix}. \tag{7}$$

Physicists recognize $mgr\sin(\varphi)$ as the magnitude of the *torque*.

The spherical pendulum's (generalized) momentum is

$$\begin{aligned} \mathsf{p} &= \mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}} \\ &= \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\sin(\varphi) & 0 \\ r\cos(\theta)\cos(\varphi) & r\sin(\theta)\cos(\varphi) & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} \dot\theta \\ \dot\varphi \end{bmatrix} \\ &= \begin{bmatrix} mr^2\sin^2(\varphi)\dot\theta \\ mr^2\dot\varphi \end{bmatrix}. \end{aligned} \tag{8}$$

The last term in Equation (6) involves the time derivative of S,

$$\dot{\mathsf{S}} = \begin{bmatrix} -r\cos(\theta)\dot\theta\sin(\varphi) - r\sin(\theta)\cos(\varphi)\dot\varphi & -r\sin(\theta)\dot\theta\cos(\varphi) - r\cos(\theta)\sin(\varphi)\dot\varphi \\ -r\sin(\theta)\dot\theta\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot\varphi & r\cos(\theta)\dot\theta\cos(\varphi) - r\sin(\theta)\sin(\varphi)\dot\varphi \\ 0 & -r\cos(\varphi)\dot\varphi \end{bmatrix}. \tag{9}$$

The last term is therefore

$$\dot{\mathsf{S}}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}} = \begin{bmatrix} mr^2\sin(\varphi)\cos(\varphi)\dot\varphi & -mr^2\sin(\varphi)\cos(\varphi)\dot\theta \\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta & 0 \end{bmatrix} \begin{bmatrix} \dot\theta \\ \dot\varphi \end{bmatrix} = \begin{bmatrix} 0 \\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 \end{bmatrix}. \tag{10}$$

Equations (7), (8), and (10) fit into Equation (6)'s framework to give the equations of motion

$$\begin{bmatrix} 0 \\ mgr\sin(\varphi) \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} mr^2\sin^2(\varphi)\dot\theta \\ mr^2\dot\varphi \end{bmatrix} - \begin{bmatrix} 0 \\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 \end{bmatrix}. \tag{11}$$

The point is that the framework provides equations of motion from which further analysis begins. Lagrange simplifies the framework by recognizing each of the three vectors in Equation (11) as a gradient.

Gradients: We have already seen that the constant force of gravity (near the surface of the earth) is a gradient of the gravitational potential $\Phi(x, y, z) = mgz$. All constant forces are gradients, of course:

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{\partial(c_1x_1+c_2x_2+\cdots)}{\partial x_1} \\ \frac{\partial(c_1x_1+c_2x_2+\cdots)}{\partial x_2} \\ \vdots \end{bmatrix} = \nabla_{\mathsf{x}} \underbrace{(c_1x_1 + c_2x_2 + \cdots)}_{\Phi(\mathsf{x})}.$$

In the language of matrix multiplication,

$$\nabla_{\mathsf{x}}\left(\mathsf{c}^T\mathsf{x}\right) = \mathsf{c},$$

which is just the multidimensional version of $\frac{d}{dx}mx = m$.

It will be convenient to write $\Phi(\mathsf{x})$ instead of $\Phi(x_1, x_2, \cdots)$, and to use vectors as the subscript in ∇_{x}. Lagrangian mechanics uses two gradients: $\nabla_{\mathsf{x}}\Phi(\mathsf{x},\mathsf{y})$ and $\nabla_{\mathsf{y}}\Phi(\mathsf{x},\mathsf{y})$. The first means "the gradient with respect to the x variables, holding the y variables constant", and the second is similarly defined.

The important gradient for Lagrangian mechanics is the gradient of a quadratic potential. Quadratic functions appear in linear algebra in second-derivative tests for extrema, and in least-squares problems.

Example: If A is a symmetric matrix, constant in x, then

$$\Phi(\mathsf{x}) = \frac{1}{2}\mathsf{x}^T\mathsf{A}\mathsf{x} = \frac{1}{2}\sum_{i,j} x_i A_{ij} x_j$$

is a *quadratic form.* Its partial derivative with respect to x_1 is, by the product rule,

$$\frac{\partial\Phi(\mathsf{x})}{\partial x_1} = \frac{1}{2}\left(\sum_j A_{1j}x_j + \sum_i x_i A_{i1}\right)$$

The first sum is the first row of $\mathsf{A}\mathsf{x}$ and the second is the first row of $\mathsf{A}^T\mathsf{x}$. Since $\mathsf{A}^T = \mathsf{A}$, the partial derivative is just the first row of $\mathsf{A}\mathsf{x}$. The gradient assembles all the partial derivatives in a vector, so

$$\nabla_{\mathsf{x}}\left(\frac{1}{2}\mathsf{x}^T\mathsf{A}\mathsf{x}\right) = \mathsf{A}\mathsf{x}. \tag{12}$$

The point is that $\mathsf{A}\mathsf{x}$ is the gradient with respect to x of the quadratic potential $\frac{1}{2}\mathsf{x}^T\mathsf{A}\mathsf{x}$ whenever A is symmetric. This is simply the multi-dimensional version of

$$\frac{d}{dx}\left(\frac{1}{2}mx^2\right) = mx.$$

Remark: The gradient and the derivative are transposes of each other:

$$\mathbf{D}_{\mathsf{x}}\Phi(\mathsf{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1}\Phi(\mathsf{x}) & \frac{\partial}{\partial x_2}\Phi(\mathsf{x}) & \cdots & \frac{\partial}{\partial x_n}\Phi(\mathsf{x}) \end{bmatrix} = \Big(\nabla_{\mathsf{x}}\Phi(\mathsf{x})\Big)^T,$$

and they serve different purposes. The derivative must satisfy the chain rule

$$\frac{d}{dt}\Phi(\mathsf{x}(t)) = \mathbf{D}_{\mathsf{x}}\Phi(\mathsf{x}(t))\frac{d}{dt}\mathsf{x}(t) = \mathbf{D}_{\mathsf{x}}\Phi(\mathsf{x}(t))\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \vdots\end{bmatrix},$$

so $\mathbf{D}_{\mathsf{x}}\Phi(\mathsf{x})$ must be a row vector. The gradient is a direction in the system space, so it must be a column vector.

Remark: The chain rule for gradients is awkward, and not simply because transposes reverse the order of multiplication. If $\mathsf{x} = \mathsf{G}(\mathsf{q})$, then the chain rule for derivatives says

$$\mathbf{D}_{\mathsf{q}}\Phi(\mathsf{G}(\mathsf{q})) = \mathbf{D}_{\mathsf{x}}\Phi(\mathsf{x})\,\mathbf{D}_{\mathsf{q}}\mathsf{G}(\mathsf{q}).$$

Take transposes of both sides for the chain rule for gradients:

$$\nabla_{\mathsf{q}}\Phi\left(\mathsf{G}(\mathsf{q})\right) = \mathbf{D}_{\mathsf{q}}\mathsf{G}(\mathsf{q})^T\,\nabla_{\mathsf{x}}\Phi(\mathsf{x}).$$

Note that the matrix $\mathbf{D}_{\mathsf{q}}\mathsf{G}(\mathsf{q})^T$ is not a gradient because gradients are vectors. In other words, while the chain rule for derivatives involves only derivative matrices, the chain rule for gradients involves both (vector) gradients and (matrix) derivatives. This technical detail suggests a strategy for proving theorems about gradients: prove the corresponding theorem about derivatives, then take transposes.

Remark: The derivative version of the gradient of a quadratic form is

$$\mathbf{D}_{\mathsf{x}}\left(\frac{1}{2}\mathsf{x}^T\mathsf{A}\mathsf{x}\right) = \mathsf{x}^T\mathsf{A}, \tag{13}$$

provided A is symmetric and constant in x. In this case, we've reversed the strategy, taking the transpose of the gradient in Equation (12) to get the derivative.

Lagrangian Framework: Our goal is to identify each of the three vectors $\mathsf{S}^T\mathsf{f}_a$, $\mathsf{p} = \mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}$, and $\dot{\mathsf{S}}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}$ from Equation (6) as a gradients.

Consider first the generalized momentum

$$\mathsf{p} = \mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}.$$

Since M is symmetric, so is $\mathsf{A} = \mathsf{S}^T\mathsf{M}\mathsf{S}$. If $\mathsf{x} = \dot{\mathsf{q}}$, then by Equation (12), the anti-gradient of $\mathsf{A}\mathsf{x}$ is

$$\underbrace{\mathsf{S}^T\mathsf{M}\mathsf{S}}_{\mathsf{A}}\underbrace{\dot{\mathsf{q}}}_{\mathsf{x}} = \nabla_{\dot{\mathsf{q}}}\left(\frac{1}{2}\dot{\mathsf{q}}^T\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}\right),$$

provided A is constant in x. But $\mathsf{A} = \mathsf{S}^T(\mathsf{q})\mathsf{M}\mathsf{S}(\mathsf{q})$ is constant in $\dot{\mathsf{q}}$ provided M is. Indeed, the point of factoring the chain rule $\frac{d}{dt}\mathsf{x}(\mathsf{G}(\mathsf{q})) = \mathsf{S}(\mathsf{q})\dot{\mathsf{q}}$ was to separate the "q"s from the "$\dot{\mathsf{q}}$"s. We therefore have

Theorem: If the (symmetric) mass matrix M is independent of the generalized velocities $\dot{\mathsf{q}}$, then

$$\mathsf{p} = \mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}} = \nabla_{\dot{\mathsf{q}}}\underbrace{\frac{1}{2}\dot{\mathsf{q}}^T\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}}_{T(\mathsf{q},\dot{\mathsf{q}})}.$$

In this differentiation, all components of q without "dots" are held constant.

Remark: The quadratic form $T = T(\mathsf{q}, \dot{\mathsf{q}})$ is the *kinetic energy* of the system. Since $\mathsf{v} = \mathsf{S}\dot{\mathsf{q}}$,

$$T = \frac{1}{2}\mathsf{v}^T\mathsf{M}\mathsf{v}.$$

In English: The kinetic energy is "one-half m v squared".

Example (spherical pendulum, continued): Equation (8) says the spherical pendulum's generalized momentum $\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}} = \begin{bmatrix} mr^2\sin^2(\varphi)\dot{\theta} \\ mr^2\dot{\varphi} \end{bmatrix}$. The kinetic energy of the spherical pendulum is therefore

$$T = \frac{1}{2}\dot{\mathsf{q}}^T\left(\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}\right) = \frac{1}{2}\left[\dot{\theta} \quad \dot{\varphi}\right]\begin{bmatrix} mr^2\sin^2(\varphi)\dot{\theta} \\ mr^2\dot{\varphi} \end{bmatrix} = \frac{1}{2}\left(mr^2\sin^2(\varphi)\dot{\theta}^2 + mr^2\dot{\varphi}^2\right).$$

The gradient with respect to $\dot{\mathsf{q}} = \begin{bmatrix} \dot{\theta} \\ \dot{\varphi} \end{bmatrix}$, treating φ (and θ) as constant, is indeed

$$\nabla_{\dot{\mathsf{q}}}T = \begin{bmatrix} \frac{\partial}{\partial\dot{\theta}}\frac{1}{2}\left(mr^2\sin^2(\varphi)\dot{\theta}^2 + mr^2\dot{\varphi}^2\right) \\ \frac{\partial}{\partial\dot{\varphi}}\frac{1}{2}\left(mr^2\sin^2(\varphi)\dot{\theta}^2 + mr^2\dot{\varphi}^2\right) \end{bmatrix} = \begin{bmatrix} mr^2\sin^2(\varphi)\dot{\theta} \\ mr^2\dot{\varphi} \end{bmatrix} = \mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}.$$

Remark: The derivation and the example illustrate the formal computation of the gradient $\nabla_{\dot{\mathsf{q}}}$ holding all the "undotted" components of q constant. This manipulation is the source of endless confusion, but the derivation exposes the correct computation: Hold all "variables" without "dots" constant and differentiate with respect to the "dotted" variables.

Remark: The physicist will very likely "see" the kinetic energy rather than "compute" it. The component of the velocity in the φ-direction is $r\dot{\varphi}$, and the component in the θ-direction is $r\sin(\varphi)\dot{\theta}$. The two directions are orthogonal, so Pythagoras's theorem says

$$\frac{1}{2}m\|\mathsf{v}\|^2 = \frac{1}{2}m\left(r^2\sin(\varphi)^2\dot{\theta}^2 + r^2\dot{\varphi}^2\right),$$

without explicitly computing S. This "shortcut" is one of the ways the Lagrangian formulation simplifies the Newtonian framework.

Once we've committed to looking for gradients, it is easy to "see" that $\mathsf{S}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}$ is the gradient with respect to $\dot{\mathsf{q}}$ of the quadratic form $T(\mathsf{q},\dot{\mathsf{q}})$. The next gradient in Lagrange's formulation, $\dot{\mathsf{S}}^T\mathsf{M}\mathsf{S}\dot{\mathsf{q}}$, is beyond my ability to "see". If its anti-gradient is obvious to you, please show me how to think about it. Perhaps the best starting point is to apply the strategy mentioned earlier: take transposes and ask

$$\text{is} \quad \dot{\mathsf{q}}^T\mathsf{S}^T\mathsf{M}\dot{\mathsf{S}} \quad \text{a derivative of some scalar?}$$

The crux of the computation is $\dot{\mathsf{S}}$, which we compute column-by-column.

$$\text{Column } j \text{ of} \quad \mathsf{S} = \frac{\partial}{\partial q_j}\mathsf{G}(\mathsf{q}) = \frac{\partial}{\partial q_j}\mathsf{x},$$

so, by the chain rule,

$$\begin{aligned}\text{column } j \text{ of } \dot{\mathsf{S}} = \frac{d}{dt}\frac{\partial}{\partial q_j}\mathsf{x} &= \frac{\partial}{\partial q_1}\left(\frac{\partial}{\partial q_j}\mathsf{x}\right)\dot{q}_1 + \cdots + \frac{\partial}{\partial q_n}\left(\frac{\partial}{\partial q_j}\mathsf{x}\right)\dot{q}_n \\ &= \frac{\partial}{\partial q_1}\frac{\partial}{\partial q_j}\left(\mathsf{x}\dot{q}_1\right) + \cdots + \frac{\partial}{\partial q_n}\frac{\partial}{\partial q_j}\left(\mathsf{x}\dot{q}_n\right),\end{aligned}$$

provided the $\dot{q}_i$ are constant under differentiation with respect to the q_k. By the equality of mixed partial derivatives,

$$\text{column } j \text{ of } \dot{\mathsf{S}} = \frac{\partial}{\partial q_j}\left[\frac{\partial}{\partial q_1}\mathsf{x}\dot{q}_1 + \cdots + \frac{\partial}{\partial q_n}\mathsf{x}\dot{q}_n\right] = \frac{\partial}{\partial q_j}\left[\mathsf{S}\dot{\mathsf{q}}\right].$$

Assemble the columns of $\dot{\mathsf{S}}$ into the matrix

$$\dot{\mathsf{S}} = \mathbf{D}_{\mathsf{q}}\left[\mathsf{S}\dot{\mathsf{q}}\right] = \mathbf{D}_{\mathsf{q}}\mathsf{v}. \tag{14}$$

Example (spherical pendulum, continued) From Equation (2), the pendulum's velocity is

$$\mathsf{S}\dot{\mathsf{q}} = \mathsf{v} = \begin{bmatrix} -r\sin(\theta)\dot{\theta}\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot{\varphi} \\ r\cos(\theta)\dot{\theta}\sin(\varphi) + r\sin(\theta)\cos(\varphi)\dot{\varphi} \\ -r\sin(\varphi)\dot{\varphi} \end{bmatrix},$$

so

$$\begin{aligned}\mathbf{D}_{\mathsf{q}}\left[\mathsf{S}\dot{\mathsf{q}}\right] &= \begin{bmatrix} \frac{\partial}{\partial\theta}(-r\sin(\theta)\dot{\theta}\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot{\varphi}) & \frac{\partial}{\partial\varphi}(-r\sin(\theta)\dot{\theta}\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot{\varphi}) \\ \frac{\partial}{\partial\theta}(\ \ r\cos(\theta)\dot{\theta}\sin(\varphi) + r\sin(\theta)\cos(\varphi)\dot{\varphi}) & \frac{\partial}{\partial\varphi}(\ \ r\cos(\theta)\dot{\theta}\sin(\varphi) + r\sin(\theta)\cos(\varphi)\dot{\varphi}) \\ \frac{\partial}{\partial\theta}(-r\sin(\varphi)\dot{\varphi}) & \frac{\partial}{\partial\varphi}(-r\sin(\varphi)\dot{\varphi}) \end{bmatrix} \\ &= \begin{bmatrix} (-r\cos(\theta)\dot{\theta}\sin(\varphi) - r\sin(\theta)\cos(\varphi)\dot{\varphi}) & (-r\sin(\theta)\dot{\theta}\cos(\varphi) - r\cos(\theta)\sin(\varphi)\dot{\varphi}) \\ (-r\sin(\theta)\dot{\theta}\sin(\varphi) + r\cos(\theta)\cos(\varphi)\dot{\varphi}) & (\ \ r\cos(\theta)\dot{\theta}\cos(\varphi) - r\sin(\theta)\sin(\varphi)\dot{\varphi}) \\ 0 & -r\cos(\varphi)\dot{\varphi} \end{bmatrix},\end{aligned}$$

in agreement with the pendulum's $\dot{\mathsf{S}}$ in Equation (9).

Our search for anti-gradients therefore reduces to

$$\text{is} \quad \dot{\mathsf{q}}^T \mathsf{S}^T \mathsf{M}\, \mathbf{D}_{\mathsf{q}} \left[\mathsf{S}\dot{\mathsf{q}}\right] \quad \text{a derivative of some scalar?}$$

Yes, by the chain rule: Equation (13), the derivative of a quadratic form, says

$$\dot{\mathsf{q}}^T \mathsf{S}^T \mathsf{M} = \mathbf{D}_{\mathsf{S}\dot{\mathsf{q}}} \left(\frac{1}{2} \left[\mathsf{S}\dot{\mathsf{q}}\right]^T \mathsf{M} \left[\mathsf{S}\dot{\mathsf{q}}\right] \right),$$

so the chain rule says

$$\dot{\mathsf{q}}^T \mathsf{S}^T \mathsf{M} \mathbf{D}_{\mathsf{q}} \left[\mathsf{S}\dot{\mathsf{q}}\right] = \mathbf{D}_{\mathsf{S}\dot{\mathsf{q}}} \left(\frac{1}{2} \left[\mathsf{S}\dot{\mathsf{q}}\right]^T \mathsf{M} \left[\mathsf{S}\dot{\mathsf{q}}\right] \right) \mathbf{D}_{\mathsf{q}} \left[\mathsf{S}\dot{\mathsf{q}}\right] = \mathbf{D}_{\mathsf{q}} \left(\frac{1}{2} \dot{\mathsf{q}}^T \mathsf{S}^T \mathsf{M} \mathsf{S} \dot{\mathsf{q}} \right).$$

Take transposes for the gradient theorem:

Theorem: If (the symmetric) M is independent of q, and if S has continuous partial derivatives (so the mixed partial derivatives are equal), then

$$\dot{\mathsf{S}}^T \mathsf{M} \mathsf{S} \dot{\mathsf{q}} = \nabla_{\mathsf{q}} \left(\frac{1}{2} \dot{\mathsf{q}}^T \mathsf{S}^T \mathsf{M} \mathsf{S} \dot{\mathsf{q}} \right).$$

Remark: The computation is subtle. For example, the derivative $\mathbf{D}_{\mathsf{S}\dot{\mathsf{q}}}$ is computed with respect to the components of the vector $\mathsf{S}\dot{\mathsf{q}}$, but the expression $\mathbf{D}_{\mathsf{S}}$ makes no sense because S is a matrix, not a vector. Even if $\dot{\mathsf{q}}$ is constant with respect to differentiation, it doesn't "factor out" of the expressions.

Lagrange's gradients turn the Newtonian framework in Equation (6) into

$$\mathsf{S}^T \mathsf{f}_a = \frac{d}{dt} \nabla_{\dot{\mathsf{q}}} T(\mathsf{q}, \dot{\mathsf{q}}) - \nabla_{\mathsf{q}} T(\mathsf{q}, \dot{\mathsf{q}}). \tag{15}$$

Equation (15) is correct no matter what forces f_a are applied to the system. If, however, f_a happens to be conservative, then the framework simplifies considerably:

Theorem: If $\mathsf{f}_a = -\nabla_{\mathsf{x}} \Phi(\mathsf{x})$, then $\mathsf{S}^T \mathsf{f}_a = -\nabla_{\mathsf{q}} \Phi(\mathsf{G}(\mathsf{q}))$.

Remark: In English, the theorem says that if f_a is a gradient, then so is $\mathsf{S}^T \mathsf{f}_a$. More precisely, if f_a is a gradient with respect to x, then $\mathsf{S}^T \mathsf{f}_a$ is a gradient with respect to q.

Proof: Since $\mathsf{x} = \mathsf{G}(\mathsf{q})$, the chain rule says

$$\mathbf{D}_{\mathsf{q}} \Phi(\mathsf{G}(\mathsf{q})) = \mathbf{D}_{\mathsf{x}} \Phi(\mathsf{x}) \underbrace{\mathbf{D}_{\mathsf{q}} \mathsf{G}(\mathsf{q})}_{\mathsf{S}}.$$

Take the negative transpose of both sides to complete the proof. ////

Example: The gravitational potential $\Phi(\mathsf{x}) = mgz$ is, in spherical coordinates, $\Phi(\mathsf{G}(\theta, \varphi)) = mgr\cos(\varphi)$. The gradient with respect to the generalized coordinates q is

$$\nabla_{\mathsf{q}} \Phi(\mathsf{q}) = \begin{bmatrix} \frac{\partial}{\partial \theta} mgr\cos(\varphi) \\ \frac{\partial}{\partial \varphi} mgr\cos(\varphi) \end{bmatrix} = \begin{bmatrix} 0 \\ -mgr\sin(\varphi) \end{bmatrix},$$

in agreement with Equation (7).

Suppose we decompose the applied forces f_a into conservative and non-conservative contributions:

$$\mathsf{f}_a = -\nabla_{\mathsf{q}} V(\mathsf{q}) + \mathsf{f}_n,$$

where $V(\mathsf{q}) = \Phi(\mathsf{G}(\mathsf{q}))$ is the potential (energy) expressed in the generalized coordinates q, and f_n is whatever force is left over. According to our convention, $\nabla_{\dot{\mathsf{q}}} V(\mathsf{q}) = \mathsf{0}$ because $V(\mathsf{q})$ is independent of $\dot{\mathsf{q}}$. Consequently, Lagrange's framework, Equation (15), is

$$\mathsf{S}^T \mathsf{f}_n = \frac{d}{dt}\nabla_{\dot{\mathsf{q}}}\left(T(\mathsf{q},\dot{\mathsf{q}}) - V(\mathsf{q})\right) - \nabla_{\mathsf{q}}\left(T(\mathsf{q},\dot{\mathsf{q}}) - V(\mathsf{q})\right). \tag{16}$$

Define the *Lagrangian* as

$$L(\mathsf{q},\dot{\mathsf{q}}) = T(\mathsf{q},\dot{\mathsf{q}}) - V(\mathsf{q}). \tag{17}$$

Like the kinetic energy, the Lagrangian is a scalar function of the generalized coordinates and velocities. Lagrange's framework simplifies to

$$\mathsf{S}^T \mathsf{f}_n = \frac{d}{dt}\nabla_{\dot{\mathsf{q}}} L(\mathsf{q},\dot{\mathsf{q}}) - \nabla_{\mathsf{q}} L(\mathsf{q},\dot{\mathsf{q}}). \tag{18}$$

Examples: The following examples illustrate Lagrange's technique for some simple systems.

Example: The spherical pendulum's Lagrangian, Equation (17), is

$$L\left(\begin{bmatrix}\theta\\ \varphi\end{bmatrix}, \begin{bmatrix}\dot\theta\\ \dot\varphi\end{bmatrix}\right) = \frac{1}{2}\left(mr^2\sin^2(\varphi)\dot\theta^2 + mr^2\dot\varphi^2\right) - mgr\cos(\varphi), \tag{19}$$

so the gradients are

$$\begin{aligned}\nabla_{\mathsf{q}} L(\mathsf{q},\dot{\mathsf{q}}) &= \begin{bmatrix}\frac{\partial L}{\partial\theta}\\ \frac{\partial L}{\partial\varphi}\end{bmatrix} = \begin{bmatrix}0\\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 + mgr\sin(\varphi)\end{bmatrix}\\ \nabla_{\dot{\mathsf{q}}} L(\mathsf{q},\dot{\mathsf{q}}) &= \begin{bmatrix}\frac{\partial L}{\partial\dot\theta}\\ \frac{\partial L}{\partial\dot\varphi}\end{bmatrix} = \begin{bmatrix}mr^2\sin^2(\varphi)\dot\theta\\ mr^2\dot\varphi\end{bmatrix}\end{aligned}$$

If $\mathsf{f}_n = \mathsf{0}$ — no one is pushing on the pendulum — the equations of motion are

$$\begin{bmatrix}0\\0\end{bmatrix} = \frac{d}{dt}\begin{bmatrix}mr^2\sin^2(\varphi)\dot\theta\\ mr^2\dot\varphi\end{bmatrix} - \begin{bmatrix}0\\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 + mgr\sin(\varphi)\end{bmatrix},$$

in agreement with Equation (11). (Note that Equation (11)'s gravitational force has moved from the left side to the right side here.)

If $\mathsf{f}_n = -\mu\mathsf{v}$ is a frictional force proportional to the velocity, but in the opposite direction, then

$$\mathsf{S}^T\mathsf{f}_n = \mathsf{S}^T(-\mu\mathsf{v}) = -\mu\mathsf{S}^T\mathsf{S}\dot{\mathsf{q}} = -\mu\begin{bmatrix}r^2\sin^2(\varphi)\dot\theta\\ r^2\dot\varphi\end{bmatrix}.$$

In this case, Lagrange's formulation, Equation (18), is

$$-\mu\begin{bmatrix}r^2\sin^2(\varphi)\dot\theta\\ r^2\dot\varphi\end{bmatrix} = \frac{d}{dt}\begin{bmatrix}mr^2\sin^2(\varphi)\dot\theta\\ mr^2\dot\varphi\end{bmatrix} - \begin{bmatrix}0\\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 + mgr\sin(\varphi)\end{bmatrix}.$$

Example: A seesaw is loaded with masses m_i at (x_i, y_i). Parameterize the (2-dimensional) system using polar coordinates, measuring the angle θ from the horizontal. Take $r_i < 0$ for masses to the left of the fulcrum, and $r_i > 0$ for masses to the right. The system vector x and the system velocity $\mathsf{v} = \dot{\mathsf{x}}$ are

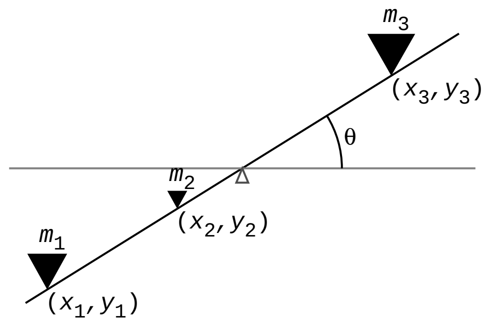

FIGURE 2: SEESAW

$$\mathsf{x} = \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{bmatrix} = \begin{bmatrix} r_1\cos(\theta) \\ r_1\sin(\theta) \\ \vdots \\ r_n\cos(\theta) \\ r_n\sin(\theta) \end{bmatrix} \quad \text{and} \quad \mathsf{v} = \underbrace{\begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}}_{\mathsf{S}} \underbrace{\begin{bmatrix} \dot\theta \end{bmatrix}}_{\dot{\mathsf{q}}}.$$

The mass matrix is diagonal, so the kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}\dot\theta \begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T \begin{bmatrix} m_1 & 0 & \cdots & 0 & 0 \\ 0 & m_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & m_n & 0 \\ 0 & 0 & \cdots & 0 & m_n \end{bmatrix} \begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix} \dot\theta \\ &= \frac{1}{2}\left(m_1 r_1^2 + \cdots + m_n r_n^2\right)\dot\theta^2. \end{aligned}$$

The physicist, of course, computes the system's kinetic energy by summing $\frac{1}{2}m_i v_i^2 = \frac{1}{2}m_i r_i^2\dot\theta^2$, without explicitly parameterizing x.

The gravitational potential energy of mass i is $m_i g y_i = m_i g r_i \sin(\theta)$. The system's gravitational potential is the sum of the masses' potentials. Check that the gravitational force acting on the system is

$$\mathsf{f}_g = \begin{bmatrix} 0 \\ -m_1 g \\ \vdots \\ 0 \\ -m_n g \end{bmatrix} = \begin{bmatrix} -\frac{\partial}{\partial x_1}\sum_i m_i g y_i \\ -\frac{\partial}{\partial y_1}\sum_i m_i g y_i \\ \vdots \\ -\frac{\partial}{\partial x_n}\sum_i m_i g y_i \\ -\frac{\partial}{\partial y_n}\sum_i m_i g y_i \end{bmatrix} = -\nabla_{\mathsf{x}} \sum_i m_i g y_i$$

and that $\mathsf{S}^T \mathsf{f}_g = \sum_i -m_i g r_i \cos(\theta) = -\nabla_\theta \sum_i m_i g r_i \sin(\theta)$. The Lagrangian for the seesaw is therefore

$$L = T - V = \left(\frac{1}{2}\sum_i m_i r_i^2\right)\dot\theta^2 - \left(\sum_i m_i g r_i\right)\sin(\theta).$$

Lagrange's equation of motion is

$$\mathsf{S}^T \mathsf{f}_n = \frac{d}{dt}\left[\left(\sum_i m_i r_i^2\right)\dot\theta\right] - \left(\sum_i m_i g r_i\right)\cos(\theta)$$

Remark: The sums are named

$$\begin{aligned} \sum_i m_i &= M = \text{the total mass or the zeroth moment} \\ \sum_i m_i r_i &= M_1 = \text{the first moment} \\ \frac{\sum_i m_i r_i}{M} &= \bar r = \text{the center of mass} \\ \sum_i m_i r_i^2 &= I = \text{the second moment or moment of inertia} \end{aligned}$$

With this notation, the equations of motion are

$$\mathsf{S}^T \mathsf{f}_n = I\ddot{\theta} - M\bar{r}\cos(\theta). \tag{20}$$

Remark: The system need not be discrete. If the seesaw has density $\rho(r)$, then the discrete sums turn into integrals:

$$M = \int \rho(r)\,dr \qquad \bar{r} = \frac{1}{M}\int r\rho(r)\,dr \qquad I = \int r^2\rho(r)\,dr.$$

With this notation, Equation (20) is still the equation of motion. In terms of the picture, the masses in Figure (2) may be spread out along the seesaw in any way we please. The seesaw itself need not be massless, and the masses need not be downward-pointing triangles. Systems in which the mass is distributed continuously rather than discretely are called *continuous media.*

Example: A double pendulum is a system with one pendulum's bob being the second pendulum's pivot. A double pendulum swinging in a plane is parameterized by the two angles made with the vertical:

$$\mathsf{x} = \begin{bmatrix} r_1\sin(\theta_1) \\ -r_1\cos(\theta_1) \\ r_1\sin(\theta_1) + r_2\sin(\theta_2) \\ -r_1\cos(\theta_1) - r_2\cos(\theta_2) \end{bmatrix}$$

so

$$\mathsf{v} = \underbrace{\begin{bmatrix} r_1\cos(\theta_1) & 0 \\ r_1\sin(\theta_1) & 0 \\ r_1\cos(\theta_1) & r_2\cos(\theta_2) \\ r_1\sin(\theta_1) & r_2\sin(\theta_2) \end{bmatrix}}_{\mathsf{S}} \underbrace{\begin{bmatrix} \dot\theta_1 \\ \dot\theta_2 \end{bmatrix}}_{\dot{\mathsf{q}}}.$$

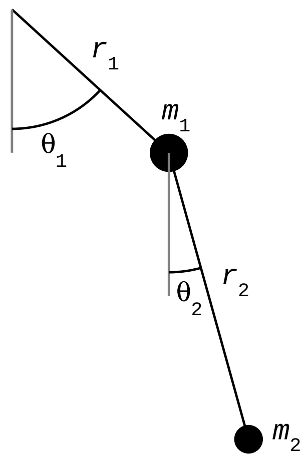

Figure 3: Double Pendulum

The mass matrix is diagonal, so the kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}\begin{bmatrix}\dot\theta_1 & \dot\theta_2\end{bmatrix} \begin{bmatrix} r_1\cos(\theta_1) & 0 \\ r_1\sin(\theta_1) & 0 \\ r_1\cos(\theta_1) & r_2\cos(\theta_2) \\ r_1\sin(\theta_1) & r_2\sin(\theta_2) \end{bmatrix}^T \begin{bmatrix} m_1 & 0 & 0 & 0 \\ 0 & m_1 & 0 & 0 \\ 0 & 0 & m_2 & 0 \\ 0 & 0 & 0 & m_2 \end{bmatrix} \begin{bmatrix} r_1\cos(\theta_1) & 0 \\ r_1\sin(\theta_1) & 0 \\ r_1\cos(\theta_1) & r_2\cos(\theta_2) \\ r_1\sin(\theta_1) & r_2\sin(\theta_2) \end{bmatrix} \begin{bmatrix} \dot\theta_1 \\ \dot\theta_2 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix}\dot\theta_1 & \dot\theta_2\end{bmatrix} \begin{bmatrix} (m_1+m_2)r_1^2 & m_2r_1r_2\cos(\theta_2-\theta_1) \\ m_2r_1r_2\cos(\theta_2-\theta_1) & m_2r_2^2 \end{bmatrix} \begin{bmatrix} \dot\theta_1 \\ \dot\theta_2 \end{bmatrix} \\ &= \frac{1}{2}\left((m_1+m_2)r_1^2\dot\theta_1^2 + 2m_2r_1r_2\cos(\theta_2-\theta_1)\dot\theta_1\dot\theta_2 + m_2r_2^2\dot\theta_2^2\right) \end{aligned}$$

The gravitational potential for the system is

$$\begin{aligned} V = m_1gy_1 + m_2gy_2 &= -m_1gr_1\cos(\theta_1) - m_2g\left(r_1\cos(\theta_1) + r_2\cos(\theta_2)\right) \\ &= -(m_1+m_2)gr_1\cos(\theta_1) - m_2gr_2\cos(\theta_2). \end{aligned}$$

If $M = m_1 + m_2$, then the Lagrangian is

$$L = T - V = \frac{1}{2}\left(Mr_1^2\dot\theta_1^2 + 2m_2r_1r_2\cos(\theta_2-\theta_1)\dot\theta_1\dot\theta_2 + m_2r_2^2\dot\theta_2^2\right) + Mgr_1\cos(\theta_1) + m_2gr_2\cos(\theta_2).$$

Lagrange's equations of motion are therefore

$$\mathsf{S}^T\mathsf{f}_n = \frac{d}{dt}\begin{bmatrix} Mr_1^2\dot\theta_1 + m_2r_1r_2\cos(\theta_2-\theta_1)\dot\theta_2 \\ m_2r_1r_2\cos(\theta_2-\theta_1)\dot\theta_1 + m_2r_2^2\dot\theta_2 \end{bmatrix} - \begin{bmatrix} m_2r_1r_2\sin(\theta_2-\theta_1)\dot\theta_1\dot\theta_2 - Mgr_1\sin(\theta_1) \\ -m_2r_1r_2\sin(\theta_2-\theta_1)\dot\theta_1\dot\theta_2 - m_2gr_2\sin(\theta_2) \end{bmatrix}.$$

Addenda:

Conservation Laws: Some conservation laws follow immediately from the Lagrangian framework. For example, the top row of Equation (11) reads

$$0 = \frac{d}{dt}\left(mr^2 \sin^2(\varphi)\dot{\theta}\right) - 0,$$

so the quantity in parentheses is constant:

$$mr^2 \sin^2(\varphi)\dot{\theta} = \text{constant}.$$

The physicist recognizes the left side as the angular momentum, so the spherical pendulum obeys the law of conservation of angular momentum whenever no non-conservative forces push it.

In general, a (generalized) coordinate q_i is *ignorable* or *cyclic* if the Lagrangian $L(\mathsf{q}, \dot{\mathsf{q}})$ is independent of q_i. In this language, θ is an ignorable coordinate of the spherical pendulum's Lagrangian, Equation (19). If q_i is ignorable, then row i of the equations of motion are

$$\left(\mathsf{S}^T \mathsf{f}_n\right)_i = \frac{d}{dt}\frac{\partial L(\mathsf{q}, \dot{\mathsf{q}})}{\partial \dot{\mathsf{q}}_i} - \underbrace{\frac{\partial L(\mathsf{q}, \dot{\mathsf{q}})}{\partial \mathsf{q}_i}}_{=0}.$$

If, in addition, $\left(\mathsf{S}^T \mathsf{f}_n\right)_i = 0$, then the system obeys the conservation law

$$\frac{\partial L(\mathsf{q}, \dot{\mathsf{q}})}{\partial \dot{\mathsf{q}}_i} = \text{constant}.$$

Moving Coordinates: The Lagrangian framework applies even if the coordinate system is moving. For example, a laboratory on the surface of the earth may feel stationary, but the earth is rotating.

If system vector

$$\mathsf{x} = \mathsf{G}(\mathsf{q}, t)$$

depends (explicitly) on time, then by the chain rule,

$$\mathsf{v} = \dot{\mathsf{x}} = \underbrace{\mathbf{D}_{\mathsf{q}} \mathsf{G}(\mathsf{q}, t)}_{\mathsf{S}(\mathsf{q},t)} \dot{\mathsf{q}} + \underbrace{\mathbf{D}_t \mathsf{G}(\mathsf{q}, t)}_{\mathsf{s}(\mathsf{q},t)} = \mathsf{S}\dot{\mathsf{q}} + \mathsf{s}.$$

In this notation, the uppercase S is a matrix; the lowercase s is a vector. Furthermore, both S and s are independent of $\dot{\mathsf{q}}$, although both may depend on q and t.

Newton's second law, Equation (1) (with the mass matrix M replacing the particle mass m), is

$$\mathsf{f} = \frac{d}{dt}\left(\mathsf{M}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right)\right).$$

The only physics argument in the derivation is how to eliminate forces of constraint. The correct technique is to left multiply both sides by S^T. Use the product rule to find the Newtonian framework

$$\mathsf{S}^T \mathsf{f}_a = \frac{d}{dt}\left(\mathsf{S}^T \mathsf{M}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right)\right) - \dot{\mathsf{S}}^T\left(\mathsf{M}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right)\right).$$

The kinetic energy is

$$T(\mathsf{q}, \dot{\mathsf{q}}, t) = \frac{1}{2}\mathsf{v}^T \mathsf{M} \mathsf{v} = \frac{1}{2}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right)^T \mathsf{M}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right).$$

The part of Equation (14) that reads $\dot{\mathsf{S}} = \mathbf{D}_{\mathsf{q}} \mathsf{v}$ is still true (exercise for the reader), but v has changed, so

$$\dot{\mathsf{S}} = \mathbf{D}_{\mathsf{q}} \mathsf{v} = \mathbf{D}_{\mathsf{q}}\left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right).$$

All that remains is to check that, by the (gradient) chain and product rules,

$$\begin{aligned} \nabla_{\mathsf{q}} T(\mathsf{q}, \dot{\mathsf{q}}, t) &= \dot{\mathsf{S}}^T \mathsf{M} \left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right), \qquad \text{and} \\ \nabla_{\dot{\mathsf{q}}} T(\mathsf{q}, \dot{\mathsf{q}}, t) &= \mathsf{S}^T \mathsf{M} \left(\mathsf{S}\dot{\mathsf{q}} + \mathsf{s}\right). \end{aligned}$$

Consequently, Lagrange's framework, Equation (16), remains valid even if $\mathsf{G}(\mathsf{q}, t)$ depends explicitly on t.

Example (rotating earth): Let x_l be coordinates in a laboratory fixed on the surface of the rotating earth. Then the coordinates in space (ignoring the motion relative to the sun and stars) are

$$\mathsf{x} = \underbrace{\begin{bmatrix} \cos(\omega t) & -\sin(\omega t) & 0 \\ \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathsf{R}(t)} \mathsf{x}_l.$$

where $\omega = \frac{2\pi}{24 \cdot 60 \cdot 60} s^{-1}$. The Newtonian reference frame — the coordinates in which Newton's law holds — is the frame of x (not x_l).

The velocity is, by the product rule,

$$\begin{aligned} \mathsf{v} &= \mathsf{R}(t)\dot{\mathsf{x}}_l + \dot{\mathsf{R}}(t)\mathsf{x}_l \\ &= \mathsf{R}(t)\mathsf{v}_l + \begin{bmatrix} -\sin(\omega t)\omega & -\cos(\omega t)\omega & 0 \\ \cos(\omega t)\omega & -\sin(\omega t)\omega & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathsf{x}_l \\ &= \mathsf{R}(t)\mathsf{v}_l + \underbrace{\begin{bmatrix} \cos(\omega t) & -\sin(\omega t) & 0 \\ \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathsf{R}(t)} \omega \underbrace{\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathsf{J}} \mathsf{x}_l \\ &= \mathsf{R}(t)\left(\mathsf{v}_l + \omega \mathsf{J} \mathsf{x}_l\right). \end{aligned}$$

The kinetic energy of the system is

$$T = \frac{1}{2}\left(\mathsf{v}_l + \omega\mathsf{J}\mathsf{x}_l\right)^T \mathsf{R}(t)^T \mathsf{M} \mathsf{R}(t) \left(\mathsf{v}_l + \omega\mathsf{J}\mathsf{x}_l\right).$$

Since $\mathsf{M} = m\mathsf{I}$ and $\mathsf{R}(\theta)$ is orthogonal,

$$\begin{aligned} T &= \frac{1}{2} m \left(\mathsf{v}_l + \omega\mathsf{J}\mathsf{x}_l\right)^T \left(\mathsf{v}_l + \omega\mathsf{J}\mathsf{x}_l\right) \\ &= \frac{1}{2} m \left(\|\mathsf{v}_l\|^2 + 2\omega \mathsf{v}_l^T \mathsf{J}\mathsf{x}_l + \omega^2 \mathsf{x}_l^T \mathsf{J}^T \mathsf{J}\mathsf{x}_l\right). \end{aligned}$$

Take $\mathsf{q} = \mathsf{x}_l$ and, therefore, $\dot{\mathsf{q}} = \mathsf{v}_l$. Then $\mathsf{S} = \mathsf{R}(t)$ and $\mathsf{s} = \mathsf{R}(t)\omega\mathsf{J}\mathsf{x}_l$ and the equations of motion are

$$\begin{aligned} \mathsf{R}(t)^T \mathsf{f}_a &= \frac{d}{dt} m \left(\mathsf{v}_l + \omega\mathsf{J}\mathsf{x}_l\right) - m\left(-\omega\mathsf{J}\mathsf{v}_l + \omega^2 \mathsf{J}^T\mathsf{J}\mathsf{x}_l\right) \\ &= \frac{d}{dt} m\mathsf{v}_l + 2m\omega\mathsf{J}\mathsf{v}_l + \omega^2\mathsf{J}^T\mathsf{J}\mathsf{x}_l. \end{aligned}$$

Physicists write $\omega\mathsf{J} =$ "$\begin{bmatrix} 0 \\ 0 \\ \omega \end{bmatrix} \times$", the cross product with the vector $\begin{bmatrix} 0 \\ 0 \\ \omega \end{bmatrix}$, the vector angular velocity. The first term on the right is the rate of change of momentum in the laboratory coordinates. The second term on the right is the *Coriolis* acceleration, and the third is the *centripetal* acceleration. It is straightforward to derive the same equations of motion from the Newtonian framework, but the point is that Lagrange's formulation is valid in moving coordinates, too.

Note that $\omega \approx 7 \times 10^{-5} s^{-1}$ — it's no wonder we don't feel the earth rotating.