From Newton to Variational Methods

Purpose: Classical mechanics is the study of Newton's three laws of motion. The first and third laws describe how forces interact. Newton's second law[1],

$$\mathbf{F} = \frac{d}{dt}\Big(m\mathbf{v}\Big), \tag{N_2}$$

quantifies how a force $\mathbf{F}$ acting on a mass m changes the momentum $\mathbf{p} = m\mathbf{v}$, where $\mathbf{v} = \frac{d}{dt}\mathbf{x}$ is the mass's velocity.

Classical mechanics is a mature theory with deep results. Advanced texts advocate using *variational methods* to solve problems. Variational methods have a different "flavor" than Newton's mechanics. While Newton speaks of forces and momenta and derivatives with respect time, variational methods speak of functionals and derivatives with respect to functions. Mechanics texts expend a good deal of effort explaining how to use variational methods, much less effort explaining why they work, and no effort at all explaining how variational methods follow from Newton's laws.

These notes describe a mathematical framework — the Newtonian framework — suitable for studying a generous number problems from classical mechanics. The framework makes clear how Newton's laws imply the variational methods, but it is also a useful platform for solving problems. In practice, *determining* the equations of motion for a particular system can be every bit as challenging as *solving* them. Any useful framework must, therefore, perform two tasks:

1) Derive suitable equations of motion, and
2) Solve (or, at least, say something useful about the solutions of) the equations.

The Newtonian framework accomplishes both tasks, in a language that undergraduate mathematicians can follow. The only tools required are a bit of linear algebra, and the product and chain rules from vector calculus.

Newtonian Framework: Three examples from classical mechanics motivate the Newtonian framework: the spherical pendulum, the lever, and a bead sliding on a rotating loop of wire. The spherical pendulum is simple enough to study in detail, but it is a single mass. The lever introduces systems consisting of many masses. The bead on a rotating wire introduces time-varying coordinate systems. Each example contributes to the Newtonian framework, and the three together suffice to determine the entire framework.

Spherical Pendulum: A pendulum bob is a (point) mass swinging from a string (or a wire or a rigid rod). The path of the bob lies on a sphere of radius r. In the spherical coordinates given in every multi-variable calculus text, the bob's position is:

$$\mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} r\cos(\theta(t))\sin(\varphi(t)) \\ r\sin(\theta(t))\sin(\varphi(t)) \\ r\cos(\varphi(t)) \end{bmatrix}. \tag{SP}$$

The pivot of the pendulum is at the origin of the coordinate system. The colatitude φ — the angle made with respect to the positive z-axis — and the longitude θ are both functions of time, but r, the length of the string holding the bob, is constant. Physicists call $\mathbf{x}$ the **system** vector and the parameters $\mathbf{q} = \begin{bmatrix} \theta \\ \varphi \end{bmatrix}$ the **generalized coordinates** of the system. Geometers say the sphere is a 2-dimensional manifold and spherical coordinates provide a coordinate patch.

[1] Newton's second law is the famous $\mathbf{F} = m\mathbf{a}$ whenever the mass is constant, but Equation (N_2) is correct when the mass is not constant. That *is* rocket science: a rocket ejects mass out the back end of the engine to provide its thrust.

Remark: If the bob swings by a flexible string, then it can *conceivably* move anywhere inside the (solid) ball of radius R. If this were the case, the radius r would have to be a function of time, too, and the constraint would be $r(t) \leq R$. The spherical pendulum models the bob as if it swings from a rigid (but massless) rod, so its motion is constrained to the *surface* of the ball, and r is constant.

The **velocity** of the bob is the time derivative (denoted by the "dot")

$$\mathbf{v} = \dot{\mathbf{x}} = \frac{d}{dt}\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} -r\sin(\theta)\sin(\varphi)\dot{\theta} + r\cos(\theta)\cos(\varphi)\dot{\varphi} \\ r\cos(\theta)\sin(\varphi)\dot{\theta} + r\sin(\theta)\cos(\varphi)\dot{\varphi} \\ -r\sin(\varphi)\dot{\varphi} \end{bmatrix} = \underbrace{\begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} \dot{\theta} \\ \dot{\varphi} \end{bmatrix}}_{\dot{\mathbf{q}}}.$$

To take the derivative of the system vector $\mathbf{x}$, just differentiate each component $x(t)$, $y(t)$ and $z(t)$ using the chain and product rules. Writing the result as a matrix times a vector splits the derivative into the "shape" matrix $\mathbf{S}$ and the parameter velocity $\dot{\mathbf{q}}$. $\mathbf{S}$ contains geometric information; $\dot{\mathbf{q}}$ contains physical information. Two important points that will recur frequently:

- $\mathbf{S}$ is a *mapping* taking parameter velocities $\dot{\mathbf{q}}$ to system velocities $\mathbf{v} = \mathbf{S}\dot{\mathbf{q}}$ in the **tangent space** of the manifold. (The tangent space at any point on the sphere is the 2-dimensional plane tangent to the sphere at that point.)
- The columns of $\mathbf{S}$ provide a basis of the tangent space at each point of the manifold.

The system's **momentum** is $\mathbf{p} = m\mathbf{v}$, where m is the bob's mass. **Newton's second law** quantifies how forces change momentum:

$$\mathbf{F} = \dot{\mathbf{p}} = \frac{d}{dt}\Big(m\mathbf{v}\Big). \qquad (N_2)$$

The vector $\mathbf{F}$ is the total force acting on the bob. The two dominant forces are the gravitational force $\mathbf{F}_g$ and the tension $\mathbf{F}_c$ in the string (or rod) which constrains the bob to move only on the sphere. **Newton's first law** says that the total force is the vector sum of the individual forces[2]. If *only* gravitational and constraining forces act on the bob, then Newton's first and second laws together say

$$\mathbf{F}_g + \mathbf{F}_c = \frac{d}{dt}\Big(m\mathbf{v}\Big). \qquad (N_{1\&2})$$

This is just a second-order (in $\mathbf{x}$), ordinary differential equation. We might hope that, given some initial values $\mathbf{x}(0) = \mathbf{x}_0$ and $\mathbf{v}(0) = \mathbf{v}_0$, we can solve the differential equation for $\mathbf{x}(t)$ — but there is a wrinkle.

The gravitational force $\mathbf{F}_g$ acts downward, and is (approximated by)

$$\mathbf{F}_g = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix}.$$

Physicists call forces which come from outside the "system" — like the gravitational force — **applied** forces. These are "known" forces. By contrast, we don't know very much about $\mathbf{F}_c$, the **force of constraint**. It comes from "inside" the system and is not given — at least, not directly.

The string and the bob pull on each other in some way that keeps the bob on the sphere, but we don't know the specifics. Any framework for solving classical mechanics problems must somehow work in spite of our ignorance of the forces of constraint. **Newton's third law**[3] is just the ticket. The idea is that the

[2] The first law is often stated as if it were a corollary of the second law, but Newton did not use it that way. Newton's Corollary I — the first proof following the statement of the three laws — is the logical first law: *A body by two forces conjoined will describe the diagonal of a parallelogram, in the same time that it would describe the sides, by those forces apart.* (See Isaac Newton, *The Principia*, translated by Andrew Motte, Prometheus Books, New York, 1995, p. 20.) Newton argues that the corollary follows from the first law; we take the corollary as the law itself.

[3] "To every action there is an equal and opposite reaction" is the usual statement.

string pulls on the bob, so the bob pulls in the opposite direction on the string. Consequently, the *direction* of the force acting on the bob is *up the string.* The constraining force is therefore parallel to $\mathbf{x}$, so it is of the form

$$\mathbf{F}_c = f(\theta, \varphi, t, \ldots)\mathbf{x} = f(\theta, \varphi, t, \ldots) \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix}.$$

The "..." in the argument of f represents variables that we don't know about but which may affect the constraining force on the bob.

The useful fact about the constraining force is that $\mathbf{F}_c$ is **orthogonal to the tangent space**. In matrix notation, this means

$$\mathbf{S}^T\mathbf{F}_c = \mathbf{0} \qquad (PVW)$$

because the columns of $\mathbf{S}$ comprise a basis of the tangent space, and $\mathbf{S}^T\mathbf{F}_c = \mathbf{0}$ means every column of $\mathbf{S}$ is orthogonal to $\mathbf{F}_c$. Check that, indeed,

$$\begin{aligned} \mathbf{S}^T\mathbf{F}_c &= \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}^T f(\ldots) \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix} \\ &= f(\ldots) \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\sin(\varphi) & 0 \\ r\cos(\theta)\cos(\varphi) & r\sin(\theta)\cos(\varphi) & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

no matter what $f(\ldots)$ is.

Remark: Condition (PVW) is called the **principle of virtual work** in some physics texts; in others, it is called an application of Newton's third law. The name "virtual work" is explained below, in the section on variational methods.

Condition (PVW) suggests a way to overcome our ignorance of $\mathbf{F}_c$: left-multiply both sides of Equation $(N_{1\&2})$ by $\mathbf{S}^T$ to eliminate the force of constraint:

$$\mathbf{S}^T(\mathbf{F}_g + \mathbf{F}_c) = \mathbf{S}^T\mathbf{F}_g = \mathbf{S}^T\frac{d}{dt}(m\mathbf{S}\dot{\mathbf{q}}).$$

The term in the middle is

$$\mathbf{S}^T\mathbf{F}_g = \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} = \begin{bmatrix} 0 \\ mgr\sin(\varphi) \end{bmatrix}.$$

The right side is more complicated. We could use the product rule to compute the derivative as written, but there is a smarter way. The product rule also says

$$\mathbf{S}^T\frac{d}{dt}(m\mathbf{S}\dot{\mathbf{q}}) = \frac{d}{dt}\left(\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}}\right) - \dot{\mathbf{S}}^T m\mathbf{S}\dot{\mathbf{q}}.$$

How do we know to use the product rule this way instead of directly computing $\frac{d}{dt}(m\mathbf{S}\dot{\mathbf{q}})$? The symmetric form $\mathbf{S}^T m\mathbf{S}$ appears frequently in applied mathematics, and it is wise to keep an eye out for it (see Strang's text[4] on applied mathematics). $\mathbf{S}^T m\mathbf{S}$ also appears in the normal equations of weighted least-squares problems as presented in linear algebra. Furthermore, $\frac{d}{dt}\left(\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}}\right)$ is the derivative of something, suggesting the "something" is of fundamental importance. If this term replaces Newton's $\frac{d}{dt}\mathbf{p}$, then $\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}}$ might be some kind of momentum.

[4] Gilbert Strang, *Introduction to Applied Mathematics*, Wellesley Cambridge Press.

In the pendulum problem,

$$\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}} = m \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix}^T \begin{bmatrix} -r\sin(\theta)\sin(\varphi) & r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ 0 & -r\sin(\varphi) \end{bmatrix} \begin{bmatrix} \dot\theta \\ \dot\varphi \end{bmatrix}$$
$$= m \begin{bmatrix} r^2\sin^2(\varphi) & 0 \\ 0 & r^2 \end{bmatrix} \begin{bmatrix} \dot\theta \\ \dot\varphi \end{bmatrix} = \begin{bmatrix} mr^2\sin^2(\varphi)\dot\theta \\ mr^2\dot\varphi \end{bmatrix}.$$

The remaining term, $\dot{\mathbf{S}}^T m\mathbf{S}\dot{\mathbf{q}}$, is more complicated. The time derivative of $\mathbf{S}$ is

$$\dot{\mathbf{S}} = \begin{bmatrix} -r\cos(\theta)\sin(\varphi)\dot\theta - r\sin(\theta)\cos(\varphi)\dot\varphi & -r\sin(\theta)\cos(\varphi)\dot\theta - r\cos(\theta)\sin(\varphi)\dot\varphi \\ -r\sin(\theta)\sin(\varphi)\dot\theta + r\cos(\theta)\cos(\varphi)\dot\varphi & r\cos(\theta)\cos(\varphi)\dot\theta - r\sin(\theta)\sin(\varphi)\dot\varphi \\ 0 & -r\cos(\varphi)\dot\varphi \end{bmatrix},$$

so

$$\dot{\mathbf{S}}^T m\mathbf{S}\dot{\mathbf{q}} = m \begin{bmatrix} r^2\sin(\varphi)\cos(\varphi)\dot\varphi & -r^2\sin(\varphi)\cos(\varphi)\dot\theta \\ r^2\sin(\varphi)\cos(\varphi)\dot\theta & 0 \end{bmatrix} \begin{bmatrix} \dot\theta \\ \dot\varphi \end{bmatrix} = \begin{bmatrix} 0 \\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 \end{bmatrix}.$$

The second order differential equation for the spherical pendulum is, therefore,

$$\begin{bmatrix} 0 \\ mgr\sin(\varphi) \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} mr^2\sin^2(\varphi)\dot\theta \\ mr^2\dot\varphi \end{bmatrix} - \begin{bmatrix} 0 \\ mr^2\sin(\varphi)\cos(\varphi)\dot\theta^2 \end{bmatrix}. \qquad (SP)_2$$

The important point is that all the terms in $(SP)_2$ are "known". The framework has accomplished half the goal by deriving the equations of motion $(SP)_2$ in terms of the generalized coordinates.

Remark: If you choose to compare $(SP)_2$ to results in a physics text, remember that our φ is measured with respect to the positive z-axis; many physics texts measure φ with respect to the negative z-axis. Many also exchange the variable names φ and θ. Coordinate system (SP) is consistent with the system in multi-variable calculus texts.

Remark: We should pause to count: the framework can describe the bob's 3-dimensional motion with a 2-dimensional system of equations because the bob moves on the 2-dimensional sphere. We eliminated one dimension by eliminating the unknown, one-dimensional, force of constraint.

Summary: In general, given a coordinate system $\mathbf{x} = \mathbf{x}(\mathbf{q})$, compute $\mathbf{S}$, the matrix from the chain rule $\mathbf{v} = \dot{\mathbf{x}} = \mathbf{S}\dot{\mathbf{q}}$. (This matrix is denoted $\mathbf{S} = \mathbf{D_q x(q)}$, the derivative with respect to the generalized coordinates $\mathbf{q}$ of the system vector $\mathbf{x}(\mathbf{q})$.) Newton's second law is

$$\mathbf{F} = \frac{d}{dt}\Big(m\mathbf{v}\Big) = \frac{d}{dt}\Big(m\mathbf{S}\dot{\mathbf{q}}\Big)$$

Use Newton's first law to split the total force $\mathbf{F}$ into a net applied force $\mathbf{F}_a$ plus a constraining force $\mathbf{F}_c$:

$$\mathbf{F}_a + \mathbf{F}_c = \frac{d}{dt}\Big(m\mathbf{S}\dot{\mathbf{q}}\Big).$$

Left multiply both sides of by $\mathbf{S}^T$. Check that Newton's third law implies (PVW), namely, $\mathbf{S}^T\mathbf{F}_c = \mathbf{0}$. This is the only step that requires any physics. If (PVW) holds, then

$$\mathbf{S}^T(\mathbf{F}_a + \mathbf{F}_c) = \mathbf{S}^T\mathbf{F}_a = \mathbf{S}^T\frac{d}{dt}\Big(m\mathbf{S}\dot{\mathbf{q}}\Big) = \frac{d}{dt}\Big(\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}}\Big) - \dot{\mathbf{S}}^T m\mathbf{S}\dot{\mathbf{q}},$$

the last equality being the product rule. The Newtonian framework is then

$$\mathbf{S}^T\mathbf{F}_a = \frac{d}{dt}\Big(\mathbf{S}^T m\mathbf{S}\dot{\mathbf{q}}\Big) - \dot{\mathbf{S}}^T m\mathbf{S}\dot{\mathbf{q}}. \qquad (NF)_1$$

At the moment, the framework has merely produced the equations of motion of the system in terms of the generalized coordinates $\mathbf{q}$ — but that is half the battle!

The Lever: The spherical pendulum is (treated as) a point mass; our second example has many point masses. If we set up the structure correctly, then the Newtonian framework for many masses will be only a slight variation on $(NF)_1$.

Suppose a (massless) lever is loaded with point masses. Think of n riders on a (massless) seesaw. It is natural to place a polar coordinate system's origin at the seesaw's fulcrum, in which case the n point masses m_i are at positions

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} r_i \cos(\theta) \\ r_i \sin(\theta) \end{bmatrix}, \quad i = 1, \ldots, n$$

on the lever. In this parameterization, the angle θ is a function of time, but the r_i (which may be positive or negative) are constant. In physics-ese: the generalized coordinate is θ, which we put in the 1-dimensional vector $\mathbf{q} = [\,\theta\,]$.

Each of the masses satisfies Newton's laws 1 and 2 expressed in Equation $(N_{1\&2})$. The goal is to arrange the equations in some way so that the framework looks very much like Equation $(NF)_1$.

The **system** is the n-fold cartesian product $R^2 \times R^2 \times \cdots \times R^2$ of the $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$, but written with the vectors stacked vertically instead of listed horizontally:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \\ x_n \\ y_n \end{bmatrix} = \begin{bmatrix} r_1 \cos(\theta) \\ r_1 \sin(\theta) \\ r_2 \cos(\theta) \\ r_2 \sin(\theta) \\ \vdots \\ r_n \cos(\theta) \\ r_n \sin(\theta) \end{bmatrix}. \tag{$(L)_1$}$$

The **system velocity v** and **system momentum p** are the "stacked" velocities and momenta

$$\mathbf{v} = \dot{\mathbf{x}} = \underbrace{\begin{bmatrix} -r_1 \sin(\theta) \\ r_1 \cos(\theta) \\ -r_2 \sin(\theta) \\ r_2 \cos(\theta) \\ \vdots \\ -r_n \sin(\theta) \\ r_n \cos(\theta) \end{bmatrix}}_{\mathbf{S}} \underbrace{[\,\dot{\theta}\,]}_{\dot{\mathbf{q}}} \qquad \mathbf{p} = \begin{bmatrix} -m_1 r_1 \sin(\theta) \\ m_1 r_1 \cos(\theta) \\ -m_2 r_2 \sin(\theta) \\ m_2 r_2 \cos(\theta) \\ \vdots \\ -m_n r_n \sin(\theta) \\ m_n r_n \cos(\theta) \end{bmatrix} [\,\dot{\theta}\,] = \mathbf{M}\mathbf{v},$$

where $\mathbf{M}$ is a diagonal matrix with blocks of $m_i \mathbf{I} = m_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ down the diagonal. $\mathbf{M}$ is the only new element here.

$$\mathbf{p}_i = m_i \frac{d}{dt} \begin{bmatrix} x_i \\ y_i \end{bmatrix}, \quad i = 1, \ldots, n.$$

Newton's second law for the system is still $\mathbf{F} = \dot{\mathbf{p}} = \frac{d}{dt}\Big(\mathbf{M}\mathbf{v}\Big)$ — check that the "stacked" vectors in the cartesian product make the matrix representation work out nicely.

$\mathbf{F}$ represents the forces acting on each mass, "stacked" into one vector, just as the positions were. The forces consist of (at least) gravitational forces and forces of constraint. The gravitational force on each point mass points downward, so

$$\mathbf{F}_g = \begin{bmatrix} 0 \\ -m_1 g \\ 0 \\ -m_2 g \\ \vdots \\ 0 \\ -m_n g \end{bmatrix}.$$

The forces of constraint are imposed by the (massless but rigid) lever. As in the case of the spherical pendulum, we don't know the exact form of the constraints, but the physics suggests that the lever can only push and pull masses in the direction it points — parallel to $\mathbf{x}$. Check that $\mathbf{x}$ is orthogonal to (the only column of) $\mathbf{S}$:

$$\mathbf{S}^T\mathbf{x} = \begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T \begin{bmatrix} r_1\cos(\theta) \\ r_1\sin(\theta) \\ \vdots \\ r_n\cos(\theta) \\ r_n\sin(\theta) \end{bmatrix} = 0.$$

As long as the constraining forces are parallel to the lever, condition (PVW) holds.

The rest of the derivation of the Newtonian framework is the identical to the derivation for the single mass: Left-multiply every "side" of $\mathbf{F} = \frac{d}{dt}\Big(\mathbf{M}\mathbf{v}\Big) = \frac{d}{dt}\Big(\mathbf{M}\mathbf{S}\dot{\mathbf{q}}\Big)$ by $\mathbf{S}^T$ to eliminate the (unknown) forces of constraint, and use the product rule to derive

$$\mathbf{S}^T\mathbf{F}_a = \frac{d}{dt}\Big(\mathbf{S}^T\mathbf{M}\mathbf{S}\dot{\mathbf{q}}\Big) - \dot{\mathbf{S}}^T\mathbf{M}\mathbf{S}\dot{\mathbf{q}}. \tag{$(NF)_2$}$$

The framework is valid whenever (PVW) holds.

The left side of Equation $(NF)_2$ for the lever, where the applied force is the (stacked) gravitational force, is

$$\mathbf{S}^T\mathbf{F}_g = \begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T \begin{bmatrix} 0 \\ -m_1 g \\ \vdots \\ 0 \\ -m_n g \end{bmatrix} = -g\sum_{i=1}^n m_i r_i\cos(\theta) = -gM\bar{r}\cos(\theta),$$

where $M = \sum_i m_i$ is the **total mass** of the system, and $\bar{r} = \sum_i \frac{m_i}{M} r_i$ is the radius of the **center of mass**

$$\begin{bmatrix} \bar{r}\cos(\theta) \\ \bar{r}\sin(\theta) \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}$$

of the system. The right side of Equation $(NF)_2$ is

$$\begin{aligned}
&\frac{d}{dt}\left(\begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T \begin{bmatrix} -m_1r_1\sin(\theta) \\ m_1r_1\cos(\theta) \\ \vdots \\ -m_nr_n\sin(\theta) \\ m_nr_n\cos(\theta) \end{bmatrix}\dot{\theta}\right) - \left(\frac{d}{dt}\begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T\right)\begin{bmatrix} -m_1r_1\sin(\theta) \\ m_1r_1\cos(\theta) \\ \vdots \\ -m_nr_n\sin(\theta) \\ m_nr_n\cos(\theta) \end{bmatrix}\dot{\theta} \\
&= \frac{d}{dt}\Big((m_1r_1^2 + \cdots + m_nr_n^2)\dot{\theta}\Big) - \begin{bmatrix} -r_1\cos(\theta)\dot{\theta} \\ -r_1\sin(\theta)\dot{\theta} \\ \vdots \\ -r_n\cos(\theta)\dot{\theta} \\ -r_n\sin(\theta)\dot{\theta} \end{bmatrix}^T \begin{bmatrix} -m_1r_1\sin(\theta) \\ m_1r_1\cos(\theta) \\ \vdots \\ -m_nr_n\sin(\theta) \\ m_nr_n\cos(\theta) \end{bmatrix}\dot{\theta} \\
&= \sum_{i=1}^n m_i r_i^2\ddot{\theta}.
\end{aligned}$$

The sum on the right, $I = \sum_{i=1}^n m_i r_i^2$, is the **moment of inertia** of the lever. In the physicists' notation, the equations of motion reduce to

$$-Mg\bar{r}\cos(\theta) = I\ddot{\theta}. \tag{$(L)_2$}$$

The framework has, again, accomplished half the goal by deriving the equations of motion.

Bead on Rotating Wire: This example introduces moving coordinate systems and exposes the last detail we'll need to construct a framework which handles a large selection of classical mechanics problems.

A circular wire loop of radius r is spinning about its vertical diameter with angular speed $\dot{\theta} = \omega$. A bead is "strung" on the wire, and slides freely. Intuitively, the faster the wire loop spins, the higher the bead rides, approaching the "equator" of the spinning hoop.

The spherical parameterization

$$\mathbf{x} = \begin{bmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{bmatrix}$$

is still applicable, but the parameter space is one-dimensional, consisting of φ only. The variable θ is no longer a parameter we need to solve for; it is a known function of time.

The system velocity is (taking derivatives with respect to φ before those with respect to θ)

$$\mathbf{v} = \frac{d}{dt}\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} r\cos(\theta)\cos(\varphi)\dot{\varphi} - r\sin(\theta)\sin(\varphi)\dot{\theta} \\ r\sin(\theta)\cos(\varphi)\dot{\varphi} + r\cos(\theta)\sin(\varphi)\dot{\theta} \\ -r\sin(\varphi)\dot{\varphi} \end{bmatrix} = \begin{bmatrix} r\cos(\theta)\cos(\varphi) & -r\sin(\theta)\sin(\varphi) \\ r\sin(\theta)\cos(\varphi) & r\cos(\theta)\sin(\varphi) \\ -r\sin(\varphi) & 0 \end{bmatrix}\begin{bmatrix} \dot{\varphi} \\ \omega \end{bmatrix}.$$

Using ω instead of $\dot{\theta}$ on the far right emphasizes that the differential equation we seek is in φ, not in θ. The generalized coordinate is $\mathbf{q} = \varphi$, and the remaining time-dependent coordinates — which we know and do not need to solve for — are $\boldsymbol{\tau} = \theta$. The matrix structure comes in "blocks" as

$$\mathbf{v} = [\,\mathbf{S} \quad \boldsymbol{\Sigma}\,]\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\boldsymbol{\tau}} \end{bmatrix}.$$

The block $\mathbf{S}$ is still the derivative with respect to the generalized coordinates of the system vector:

$$\mathbf{S} = \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix} = \frac{\partial \mathbf{x}}{\partial \varphi} = \mathbf{D}_{\mathbf{q}}\mathbf{x}(\mathbf{q}, \boldsymbol{\tau})$$

The block $\boldsymbol{\Sigma}$ is the derivative with respect to all the remaining time-dependent variables:

$$\boldsymbol{\Sigma} = \begin{bmatrix} -r\sin(\theta)\sin(\varphi) \\ r\cos(\theta)\sin(\varphi) \\ 0 \end{bmatrix} = \frac{\partial \mathbf{x}}{\partial \theta} = \mathbf{D}_{\boldsymbol{\tau}}\mathbf{x}(\mathbf{q}, \boldsymbol{\tau}).$$

With this convention of blocking the derivative matrix into $\mathbf{S}$ and $\boldsymbol{\Sigma}$, the rest of the framework in unchanged. Newton's first and second laws are still

$$\mathbf{F}_a + \mathbf{F}_c = \frac{d}{dt}(\mathbf{M}\mathbf{v}) = \frac{d}{dt}\left(\mathbf{M}\,[\,\mathbf{S} \quad \boldsymbol{\Sigma}\,]\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\boldsymbol{\tau}} \end{bmatrix}\right).$$

The question is how to eliminate the unknown forces of constraint. Specifically, what are the forces of constraint orthogonal to? The columns of both $\mathbf{S}$ *and* $\boldsymbol{\Sigma}$? No. The force of the wire pushing the bead is not orthogonal to $\boldsymbol{\Sigma}$ (check that $\boldsymbol{\Sigma}$ points horizontally in the direction of rotation, which is the direction the wire is pushing the bead). Newton's third law says that, if the bead isn't sliding along the wire, then there must be no net force *along the wire* — in the direction of $\mathbf{S}$. We conclude that the constraining forces are orthogonal to the columns of $\mathbf{S}$ and not necessarily to the columns of $\boldsymbol{\Sigma}$. In other words, condition (PVW) is still valid. (This point is worth thinking about because the physics of the constraint determines the correct matrix blocking.) We therefore left multiply both sides by $\mathbf{S}^T$ to eliminate $\mathbf{F}_c$:

$$\begin{aligned} \mathbf{S}^T\mathbf{F}_a &= \mathbf{S}^T\frac{d}{dt}\left(\mathbf{M}\,[\,\mathbf{S} \quad \boldsymbol{\Sigma}\,]\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\boldsymbol{\tau}} \end{bmatrix}\right) \\ &= \frac{d}{dt}\left(\mathbf{S}^T\mathbf{M}\,[\,\mathbf{S} \quad \boldsymbol{\Sigma}\,]\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\boldsymbol{\tau}} \end{bmatrix}\right) - \dot{\mathbf{S}}^T\mathbf{M}\,[\,\mathbf{S} \quad \boldsymbol{\Sigma}\,]\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\boldsymbol{\tau}} \end{bmatrix}, \end{aligned} \qquad (NF)_3$$

the last equality being the product rule (again).

The applied force is the same gravitational force as acted on the pendulum bob, so the left side is

$$\mathbf{S}^T\mathbf{F}_a = \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} = mgr\sin(\varphi).$$

(The parameter space is 1-dimensional, so $\mathbf{S}^T\mathbf{F}_a$ is a scalar instead of a 2-dimensional vector as it was for the spherical pendulum.) The mass matrix is just $\mathbf{M} = m\mathbf{I}$ since the bead is a point mass. The right side is therefore

$$\begin{aligned}
&\frac{d}{dt}\left(m \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix}^T \begin{bmatrix} r\cos(\theta)\cos(\varphi) & -r\sin(\theta)\sin(\varphi) \\ r\sin(\theta)\cos(\varphi) & r\cos(\theta)\sin(\varphi) \\ -r\sin(\varphi) & 0 \end{bmatrix} \begin{bmatrix} \dot\varphi \\ \omega \end{bmatrix} \right) \\
&\quad - m \begin{bmatrix} -r\cos(\theta)\sin(\varphi)\dot\varphi - r\sin(\theta)\cos(\varphi)\omega \\ -r\sin(\theta)\sin(\varphi)\dot\varphi + r\cos(\theta)\cos(\varphi)\omega \\ -r\cos(\varphi)\dot\varphi \end{bmatrix}^T \begin{bmatrix} r\cos(\theta)\cos(\varphi) & -r\sin(\theta)\sin(\varphi) \\ r\sin(\theta)\cos(\varphi) & r\cos(\theta)\sin(\varphi) \\ -r\sin(\varphi) & 0 \end{bmatrix} \begin{bmatrix} \dot\varphi \\ \omega \end{bmatrix} \\
&= \frac{d}{dt}\left(m \begin{bmatrix} r^2 & 0 \end{bmatrix} \begin{bmatrix} \dot\varphi \\ \omega \end{bmatrix} \right) - m \begin{bmatrix} 0 & r^2\sin(\varphi)\cos(\varphi)\omega \end{bmatrix}^T \begin{bmatrix} \dot\varphi \\ \omega \end{bmatrix} \\
&= \frac{d}{dt}\left(mr^2\dot\varphi \right) - mr^2\sin(\varphi)\cos(\varphi)\omega^2.
\end{aligned}$$

The equations of motion for the bead are, therefore,

$$mgr\sin(\varphi) = mr^2\ddot\varphi - mr^2\omega^2\sin(\varphi)\cos(\varphi). \qquad (BW)_1$$

Summary: The Newtonian framework $(NF)_3$ offers a unified approach to classical mechanics problems. The examples above illustrate how to *derive* equations of motion from the framework. The framework also suggests methods to *analyze* and *solve* the equations. The simplest equations to solve are the equilibrium equations — states for which $\dot{\mathbf{q}} = \mathbf{0}$, and the masses are at rest in the coordinate system. These are the subject of the next section, along with the variational interpretation of the framework.

Statics and Variational Methods: Equilibrium is defined by $\dot{\mathbf{q}} = \mathbf{0}$. The Newtonian framework $(NF)_3$ simplifies the study of equilibrium to

$$\begin{aligned}
\mathbf{S}^T\mathbf{F}_a &= \frac{d}{dt}\left(\mathbf{S}^T\mathbf{M} \begin{bmatrix} \mathbf{S} & \boldsymbol{\Sigma} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \dot{\boldsymbol{\tau}} \end{bmatrix} \right) - \dot{\mathbf{S}}^T\mathbf{M} \begin{bmatrix} \mathbf{S} & \boldsymbol{\Sigma} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \dot{\boldsymbol{\tau}} \end{bmatrix} \\
&= \frac{d}{dt}\left(\mathbf{S}^T\mathbf{M}\boldsymbol{\Sigma}\dot{\boldsymbol{\tau}} \right) - \dot{\mathbf{S}}^T\mathbf{M}\boldsymbol{\Sigma}\dot{\boldsymbol{\tau}} \qquad (EQ)_1 \\
&= \mathbf{S}^T \frac{d}{dt}\left(\mathbf{M}\boldsymbol{\Sigma}\dot{\boldsymbol{\tau}} \right),
\end{aligned}$$

the last equality being the product rule. If the coordinate frame is not moving, then $\dot{\boldsymbol{\tau}} = \mathbf{0}$, and the equilibrium reduces to

$$\mathbf{S}^T\mathbf{F}_a = \mathbf{0} \qquad \text{if } \dot{\boldsymbol{\tau}} = \mathbf{0}.$$

Example: The spherical pendulum's equilibrium condition is

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

The pendulum is in equilibrium iff $mgr\sin(\varphi) = 0$, which means $\varphi = 0$ or π — the pendulum is pointing straight up or hanging straight down. The upward-pointing equilibrium is obviously **unstable**; the downward-hanging equilibrium is **stable** — even though we haven't defined "stable".

Example: The lever's equilibrium condition is

$$\mathbf{S}^T\mathbf{F}_g = \begin{bmatrix} -r_1\sin(\theta) \\ r_1\cos(\theta) \\ \vdots \\ -r_n\sin(\theta) \\ r_n\cos(\theta) \end{bmatrix}^T \begin{bmatrix} 0 \\ -m_1 g \\ \vdots \\ 0 \\ -m_n g \end{bmatrix} = -g \underbrace{\sum_{i=1}^{n} m_i r_i}_{M\bar{r}} \cos(\theta) = 0.$$

The lever is in equilibrium iff $\cos(\theta) = 0$ or $\bar{r} = 0$.

If $\bar{r} = 0$, then the center of mass rests directly over the fulcrum. This is the **law of the lever**

$$\sum m_i r_i = 0.$$

When this condition is met, the lever may be put in any position and, if the initial velocity $\dot{\theta} = 0$, it will stay in that position forever. (Some physics texts suggest only that, if the lever starts in the horizontal position with $\dot{\theta} = 0$, it will stay horizontal forever, but the framework gives the complete story.)

Whether or not $\bar{r} = 0$, there is an equilibrium at $\theta = \pm\frac{\pi}{2}$, where the lever points straight up or straight down. The stability of these equilibria depends on where the center of mass lies: the equilibrium is stable if the center of mass lies below the fulcrum, and unstable if the center of mass lies above the fulcrum. (We defer the proof; these are supposed to be "obvious" observations.)

The point is that the framework has revealed all the solutions, even the ones often skipped in texts.

Example: The bead on a rotating loop of wire is in equilibrium when $\dot{\mathbf{q}} = \dot{\varphi} = 0$ — when the bead is not sliding up or down the wire. The bead is still moving because $\dot{\boldsymbol{\tau}} = \omega$ is non-zero, but it is not sliding. The left side of the equilibrium condition $(EQ)_1$ is

$$\mathbf{S}^T\mathbf{F}_a = \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix} = mgr\sin(\varphi).$$

The last of the equilibrium conditions $(EQ)_1$ is

$$\begin{aligned}\mathbf{S}^T\frac{d}{dt}\left(\mathbf{M\Sigma}\dot{\boldsymbol{\tau}}\right) &= \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix}^T \frac{d}{dt}\left(m \begin{bmatrix} -r\sin(\theta)\sin(\varphi) \\ r\cos(\theta)\sin(\varphi) \\ 0 \end{bmatrix} \omega \right) \\ &= \begin{bmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ -r\sin(\varphi) \end{bmatrix}^T \left(m \begin{bmatrix} -r\sin(\theta)\cos(\varphi)\dot{\varphi} - r\cos(\theta)\sin(\varphi)\omega \\ r\cos(\theta)\cos(\varphi)\dot{\varphi} - r\sin(\theta)\sin(\varphi)\omega \\ 0 \end{bmatrix} \omega + m \begin{bmatrix} -r\sin(\theta)\sin(\varphi) \\ r\cos(\theta)\sin(\varphi) \\ 0 \end{bmatrix} \dot{\omega} \right) \\ &= -mr^2\omega^2\sin(\varphi)\cos(\varphi).\end{aligned}$$

Consequently, the equilibrium condition is

$$mgr\sin(\varphi) = -mr^2\omega^2\sin(\varphi)\cos(\varphi) \qquad \text{or} \qquad mr^2\omega^2\sin(\varphi)\left(\frac{g}{r\omega^2} + \cos(\varphi)\right) = 0.$$

The bead is in equilibrium when $\sin(\varphi) = 0$ — at the top and bottom of the hoop. Intuitively, the top is always unstable, and the bottom is unstable if ω is large enough. The third equilibrium is

$$\varphi_e = \arccos\left(-\frac{g}{r\omega^2}\right) = \pi - \arccos\left(\frac{g}{r\omega^2}\right).$$

This third equilibrium does not exist if the loop is spinning so slowly that $r\omega^2 < g$. In this case, the bead rides at the bottom of the loop. If $r\omega^2 > g$, the bottom of the loop becomes an unstable equilibrium and the bead rides higher on the loop. The larger ω^2 is, the larger φ_e is, and, therefore, the higher the bead rides. Since $\lim_{\omega\to\infty}\varphi_e = \frac{\pi}{2}$, the bead never rises above the "equator" (defined by $z = 0$). These are the behaviors predicted in the discussion of the bead above.

Note that the condition for equilibrium is independent of $\dot{\omega}$ — the wire may be accelerating, but the equilibrium position depends only on ω.

Variational Methods: The equilibrium condition $(EQ)_1$ leads to a variational statement by asking "what does it mean for a vector to be zero?". A vector $\mathbf{b} \in R^n$ is the zero vector if (and only if) all of its components are zero. To "compute" the components of $\mathbf{b}$, take the scalar product

$$b_i = \mathbf{e}_i^T\mathbf{b} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix},$$

where the unit vector $\mathbf{e}_i$ is zero in every component except the i^{th}, which is 1. When $\mathbf{b}$ is n-dimensional, the statement $\mathbf{b} = \mathbf{0}$ is therefore equivalent to

$$\mathbf{e}_i^T\mathbf{b} = 0 \qquad \forall\ i = 0, 1, \ldots, n.$$

This amounts to formally checking that the n scalar components of $\mathbf{b}$ are zero to deduce that $\mathbf{b} = \mathbf{0}$.

As far as determining whether $\mathbf{b} = \mathbf{0}$ is concerned, there is nothing special about the unit coordinate vectors $\mathbf{e}_i$. A moment's thought makes it clear that

Lemma: A vector $\mathbf{b} \in R^n$ is equal to $\mathbf{0}$ iff

$$\mathbf{w}^T\mathbf{b} = 0 \qquad \forall\ \mathbf{w} \in R^n. \tag{W}$$

(The easy proof takes $\mathbf{w} = \mathbf{b}$.)

Remark: Condition (W) is called a **weak** formulation of $\mathbf{b} = \mathbf{0}$ (for reasons that become clear in functional analysis, where dual spaces can be endowed with a weak topology).

Physicists have an interpretation of the weak formulation of the equilibrium condition $(EQ)_1$. To understand the interpretation, suppose that the coordinate system is not moving, so that condition $(EQ)_1$ says

$$\mathbf{S}^T\mathbf{F}_a = \mathbf{0}.$$

The weak formulation is

$$\mathbf{w}^T\mathbf{S}^T\mathbf{F}_a = 0 \quad \text{for all } \mathbf{w} \text{ in the parameter space.}$$

Since $\mathbf{w}^T\mathbf{S}^T = (\mathbf{S}\mathbf{w})^T$, the weak formulation is equivalent to

$$(\mathbf{S}\mathbf{w})^T\,\mathbf{F}_a = 0 \quad \text{for all } \mathbf{w} \text{ in the parameter space.}$$

Recall that $\mathbf{S}$ is a *mapping* taking parameter velocities $\dot{\mathbf{q}}$ to system velocities $\mathbf{v}$ in the tangent space. Up to now, the only parameter velocities we have considered are the velocities $\dot{\mathbf{q}}$ of the solutions we seek. The weak formulation says we need to widen our consideration to *any* parameter velocity $\mathbf{w}$. If we do so, then the weak formulation is

$$\widehat{\mathbf{v}}^T\mathbf{F}_a = 0 \quad \text{for all (possible) system velocities } \widehat{\mathbf{v}}\ (= \mathbf{S}\mathbf{w}) \text{ in the tangent space.}$$

The "hat" over the $\mathbf{v}$ reminds us that $\widehat{\mathbf{v}}$ is an arbitrary velocity, not restricted to velocities of the solutions we seek.

The physicist has a nice visual interpretation of this rule. First, replace $\widehat{\mathbf{v}}^T$ with a scalar multiple

$$\widehat{\mathbf{v}}^T \delta t = \delta \mathbf{x}.$$

Mathematically, the tangent space is invariant under scalar multiplication, so the $\delta\mathbf{x}$ span the same space as the $\widehat{\mathbf{v}}$. The physicist then interprets $\delta\mathbf{x}$ as an "infinitesimal" system increment by saying δt is "infinitesimal" and calling $\delta\mathbf{x}$ a **virtual displacement**. (Note in particular that δ is not an operator here; it is simply an indicator that $\delta\mathbf{x}$ is a vector in the tangent space.) If the manifold isn't flat, $\mathbf{x} + \delta\mathbf{x}$ will not necessarily be a point on the manifold and, therefore, not an actual displacement, but the physicist views the "infinitesimal" displacement in the manifold as being equivalent to a tangent vector.

The physicist is now prepared to interpret the weak formulation

$$(\delta\mathbf{x})^T \mathbf{F}_a = 0 \qquad \text{for all virtual displacements } \delta\mathbf{x}$$

as a statement about work. The (actual) work moving an object along a path $\mathbf{x}(t)$ using a force $\mathbf{F}$ is

$$W = \int_{t_0}^{t_1} \mathbf{x}(t)^T \mathbf{F}\, dt.$$

When $t_1 - t_0 = \delta t$ is "infinitesimal", then the work is

$$\delta W = (\delta\mathbf{x})^T \mathbf{F}.$$

To indicate that $\delta\mathbf{x}$ can come from any path at all, not merely the path of the solution to the problem, the physicist calls δW the **virtual work**. As a pedagogical tool, then, the physicist replaces a discussion of $\widehat{\mathbf{v}}$s in the tangent space with an infinitesimal argument that is familiar and easy for physics to "see". Here is a typical explanation in a physics text[5] (our $\delta\mathbf{x}$ is the "stacked" system of $\delta\mathbf{r}_i$):

> ... *freeze* the system as some instant of time t; then *imagine* the particles displaced amounts $\delta\mathbf{r}_i$ consistent with the conditions of constraint. This is called a **virtual displacement**. We use $\delta\mathbf{r}_i$ rather than $\mathbf{r}_i$ to distinguish virtual displacements from real displacements. We then apply our work idea not to real displacements but to virtual displacements ...

Freezing the system and making displacements consistent with the conditions of constraint corresponds to varying $\mathbf{q}$, making the "infinitesimal" displacements in the tangent space spanned by the columns of $\mathbf{S}$. This is why we left-multiply by $\mathbf{S}^T$ and not by $\mathbf{\Sigma}^T$. Only variations in $\mathbf{q}$ are permitted because $\boldsymbol{\tau}$ can't vary.

To conclude, the physicist views the equation of equilibrium $\mathbf{S}^T\mathbf{F}_a = \mathbf{0}$ as the statement that the virtual work along *any* virtual displacement $\delta\mathbf{x}$ is zero. Confusingly, some physics texts[6] restrict attention to systems satisfying our condition (PVW) and call the equilibrium condition $\mathbf{S}^T\mathbf{F}_a = \mathbf{0}$ the principle of virtual work. (The devilish detail is in the subscript: is the principle of virtual work a statement about forces of constraint $\mathbf{F}_c$ or about the applied forces $\mathbf{F}_a$ in equilibrium?)

The good news is that the physicists' point of view very often reduces problems to analyzing a picture.

Example: The applied force acting on the spherical pendulum is $\mathbf{F}_g$, which points vertically downward. The physicist asks at what points on the sphere is the virtual work $(\delta\mathbf{x})^T \mathbf{F}_g$ equal to zero, where $\delta\mathbf{x}$ is an (infinitesimal) virtual increment on the sphere. The mathematician asks the same question in a different language: at what points on the sphere is every $\delta\mathbf{x}$ in the tangent space orthogonal to $\mathbf{F}_g$. The answers are the same: at the top and bottom of the sphere, where $\delta\mathbf{x}$ is horizontal — the same answers we computed in analytic detail above.

Example: The applied force $\mathbf{F}_a$ acting on the lever is the gravitational force acting on each mass. The force is directed vertically downward at each mass. The physicist asks at what points in the system space is the virtual work $(\delta\mathbf{x})^T \mathbf{F}_a$ equal to zero, where $\delta\mathbf{x}$ is an (infinitesimal) virtual increment of the system of points

[5] M. G. Calkin, *Lagrangian and Hamiltonian Mechanics*, World Scientific Press, p. 31. Italics and bold face are Calkin's.

[6] Herbert Goldstein, *Classical Mechanics*, Addison-Wesley, 2nd edition, 1980, p. 17.

on the lever. Without parameterizing the system, the physicist puts the seesaw in the horizontal position, where the only allowed motions are vertical, and writes down the answer:

$$\sum_i \begin{bmatrix} 0 & \delta y_i \end{bmatrix} \begin{bmatrix} 0 \\ -m_i g \end{bmatrix} = -g \sum_i m_i r_i \delta\theta = 0,$$

the law of the lever found earlier. (Since the seesaw is horizontal, the $\delta y_i = r_i \delta\theta$ just by looking at the picture). The idea is that the physicist has found a quick and reliable (= visual) method of deriving the condition of equilibrium. The question of what happens when the seesaw is not horizontal is not difficult to analyze in the same way, but is often skipped in texts, as are the (obvious) equilibria of the vertical seesaw.

Remark: Part of the appeal of this technique is that we are permitted to ignore the forces of constraint altogether — provided we've checked that they satisfy (PVW).

Remark: Part of the appeal of this technique is that the picture is easier to analyze than the formulas. Indeed, with a picture, there is no need to parameterize the manifold at all, much less compute $\mathbf{S}$.

Summary: The "increments" $\delta\mathbf{x}$ are increments of system vectors $\mathbf{x}$, which we are thinking of as functions of time. By incrementing the *function* $\mathbf{x}$ rather than just the independent variable t, we cross the line separating ordinary calculus from the calculus of variations. We have moved from Newton to variational methods.