On Dynamics

Spring 2020

Revision: 20–Dec–2019

Contents

1 Resources

Below is an annotated list of resources for topics in dynamics. They are in order of relative importance.

Greenwood (1977)	*The* classic text on Engineering Dynamics.
Greenwood (2003)	A great companion to the classic Greenwood text. This one includes many alternative formulations of equations of motion, such as Maggi's equation and Hamilton's principle.
Meirovitch (2001)	A classic vibrations text with the classical derivation starting with Newton's 2nd Law, then progressing to D'Alembert's Principle, Extended Hamilton's Principle, and Lagrange's Equation.
Udwadia and Kalaba (2007)	The book on the Udwadia-Kalaba equation for constrained systems. The original landmark paper is Udwadia and Kalaba (1992). Another title of note is Udwadia and Schutte (2010), for its treatment of rigid bodies and finite rotations. Issues surrounding the errors that can build-up in the constraints is an active topic in the literature,a comprehensive state of the literature can be found in the review of Marques, Souto, and Flores (2016).
Roithmayr and Hodges (2016)	This is an updated version of the classic text by Kane and Levinson (1985) (which is both hard to find and hard to read). Kane's method is very useful for complicated multibody systems.
Baruh (2014)	As well as the earlier text Baruh (1999) are solid general books that cover a wide range of topics, and have some great examples. Three-dimensional rotations is nicely presented here.
Schaub and Junkins (2014)	Specifically targeted towards aerospace problems.

2 Primer

This section is a brief review of what is typically covered in sophomore/junior course in dynamics. It is not intended to be comprehensive, just a reminder.

2.1 Kinematics

This is a made up word from the Greek "to move." The study of kinematics is about describing the motion of an object, not the cause of the motion. It all revolves around the following definitions, if $\boldsymbol{r}$ is a position vector then

$$\boldsymbol{v} \coloneqq \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \tag{1}$$

$$\boldsymbol{a} \coloneqq \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}. \tag{2}$$

These deceptively simple looking relations for velocity $\boldsymbol{v}$ and acceleration $\boldsymbol{a}$ have many layers of subtly. There are times, common in machine motion, where we need to use the several higher-order derivatives[1] but we can skip that for now. The kinematics problems usually contain the more difficult mathematics encountered in dynamics problems. They contain the geometric constraints in most problems, such as the nonlinearities of rotation.

[1] These are jerk, snap, crackle, and pop. Listing them always brings a smile to my face.

2.1.1 Frames of reference

Much ado is placed on frames of reference, and rightly so, it is a really important topic.

frames of reference

2.1.2 Unit vectors

Recall that a unit vector is simply a mathematical way to express a direction. I'll use the hat notation (e.g. $\hat{\boldsymbol{u}}$) to emphasize that a vector is of unit length (and also is dimensionless). The universal property of a unit vectors is

$$\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{u}} = 1\,. \tag{3}$$

We're all likely to be very comfortable when these directions do not change over time. However, it is usually convenient to use directions that do change in time. This happens regularly when we have a sensor that is attached to a moving body.

include figure of sensor on moving body

Recall the velocity equation for the velocity of point spinning about another point

$$\boldsymbol{v}_B = \boldsymbol{v}_A + \underbrace{\boldsymbol{\omega} \times \boldsymbol{r}_{B/A}}_{\text{B rotating wrt A}} \tag{4}$$

which says that the velocity of point B is the velocity of point A plus the relative velocity of B with respect to A. This equation is also useful for computing the derivative of unit vectors. Let's consider a unit vector $\hat{\boldsymbol{u}}$ that is spinning about with angular velocity $\boldsymbol{\Omega}$, then we have

$$\frac{\mathrm{d}\hat{\boldsymbol{u}}}{\mathrm{d}t} = \underbrace{\frac{\partial \hat{\boldsymbol{u}}}{\partial t}}_{\text{change of length}} + \underbrace{\boldsymbol{\Omega} \times \hat{\boldsymbol{u}}}_{\text{change of direction}} \tag{5}$$

Since the length of the unit vector is fixed by definition, then that term is zero leaving only the second term. This gives the very useful relation

$$\frac{\mathrm{d}\hat{\boldsymbol{u}}}{\mathrm{d}t} = \boldsymbol{\Omega} \times \hat{\boldsymbol{u}} \tag{6}$$

for all unit vectors. With this expression, we can now treat complicated problems in any moving (specifically rotating) frame of reference[2].

2.1.3 Common frames of reference

The Cartesian frame is the mutually orthogonal unit-vector triad that is usually denoted by $\{\hat{\boldsymbol{i}}\hat{\boldsymbol{j}}\hat{\boldsymbol{k}}\}$. Sometimes, capital letters are used to denote this frame when it is not rotating while lower case letters are used for a moving frame. There's no single standard or agreement on a convention.

After Cartesian coordinates we will likely encounter: polar (2D), cylindrical (3D), and normal–tangential[3]. Other systems exist, such as spherical and general curvilinear but we will likely not meet them here. There is no single convention for the label associated with each frame, but it's usually clear by the use of subscripts.

As an example, we can write a position vector as

$$\boldsymbol{r} = x(t)\hat{\boldsymbol{i}} + y(t)\hat{\boldsymbol{j}} + z(t)\hat{\boldsymbol{k}} = r\hat{\boldsymbol{e}}_r$$

The velocity is also tangent to the path

include figure here

$$\boldsymbol{v} = v\hat{\boldsymbol{e}}_t \tag{7}$$

[2] Provided enough time and snacks are provided.
[3] Also know as the Frenet–Serret frame.

where the unit tangent is

$$\hat{\boldsymbol{e}}_t = \frac{\mathrm{d}\boldsymbol{r}(s)}{\mathrm{d}s} = \frac{\mathrm{d}\boldsymbol{r}(t)}{\mathrm{d}t}\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^{-1} = \frac{\boldsymbol{v}}{|\boldsymbol{v}|} \tag{8}$$

with the arc length s.

2.1.4 Vectors are not columns of a matrix

A vector quantity is a magnitude and a direction, and since we work in either 2D or 3D it is really useful to write down components of the magnitude in terms of unit vectors. This is great. It is also common in computing that we can use a column to *represent* a vector. We just order the coefficients and list them in a tuple.

$$3\hat{\boldsymbol{i}} - 2\hat{\boldsymbol{k}} \to \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}$$

The word *represent* here has some very deep mathematical meaning. To us, it means that the column lists the components in a particular set of directions (coordinate frame). We have to remember what those directions are, what units they have, and if they are moving.

When we use matrix notation[4] we have to be extremely careful for at least two reasons.

1. Mixing frames. We could inadvertently add two columns that were originally defined in different coordinate frames. This mistake is very easy to make, since no frame information is conveyed in the matrix.

2. Derivatives. We need to remember if we are in a moving frame, as the column will not remind us. Imagine we have a vector in polar coordinates $r(t)$ and $\theta(t)$

$$\boldsymbol{r}(t) = r(t)\boldsymbol{e}_r$$

and we write it as a column[5]

$$\boldsymbol{r}(t) \to \mathsf{\mathbf{r}}(t) = \begin{bmatrix} r(t) \\ 0 \end{bmatrix}$$

When we compute the velocity

$$\begin{aligned} \boldsymbol{v}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}(t) &= \dot{r}(t)\hat{\boldsymbol{e}}_r + r(t)\dot{\hat{\boldsymbol{e}}}_r \\ &= \dot{r}(t)\hat{\boldsymbol{e}}_r + r(t)\left(\dot{\theta}\hat{\boldsymbol{e}}_z \times \hat{\boldsymbol{e}}_r\right) \\ &= \dot{r}(t)\hat{\boldsymbol{e}}_r + r(t)\dot{\theta}(t)\hat{\boldsymbol{e}}_\theta \end{aligned}$$

We get the expected results, but when we differentiate the column matrix directly

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} r(t) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{r}(t) \\ 0 \\ 0 \end{bmatrix}$$

[4]Since we are basically just dealing with lists of components it is alternatively called tuple notation.

[5]Note that I write the column in a bold sans serif font $\mathsf{\mathbf{r}}(t)$, and the vector is in italics $\boldsymbol{r}(t)$. I prefer this notation as it cleanly indicates what type of quantity we have. It also permits us to write the same quantity as a vetor or a tuple.

and in vector notation that would be $\dot{r}(t)\hat{\boldsymbol{e}}_r$. This is missing the term that comes from the moving frame: $r\dot{\theta}\hat{\boldsymbol{e}}_\theta$. However we can use our equation for the derivative of a rotating vector here

$$\frac{{}^N\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{{}^B\mathrm{d}\mathbf{r}}{\mathrm{d}t} + {}^{B/N}\boldsymbol{\omega} \times \mathbf{r}$$
$$= \begin{bmatrix} \dot{r} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dot{\theta} \end{bmatrix} \times \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{r} \\ r\dot{\theta} \\ 0 \end{bmatrix}$$

and we get the same result as before. The extra superscripts are one of the conventions to indicate which frame the tuple is in: B for body-fixed (rotating), and N for Newtonian (inertial).

The big takeaway is that vectors and tuples (column vectors, matrices, arrays, etc) are different things but we can use either notation as long as we are very careful.

2.1.5 Example in multiple frames

Add examples of switching between multiple frames

2.2 Kinetics

Add section on $\boldsymbol{F} = m\boldsymbol{a}$

3 Equations of motion beyond Newton

3.1 Coordinates

Let's say we want to describe the motion of a particle, or the location of a point on a body. If we are in a fixed Cartesian coordinates we could write the position as

$$\boldsymbol{r} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}} \tag{9}$$

where x, y, z are physical coordinates. We have likely seen that it can be advantageous to use other variables in a problem to make life easier. Consider the simple pendulum: we do not use x, y, instead we use θ. We can build the position using a transformation like $x = L\sin\theta$, $y = -\cos\theta$, but we need only a single coordinate to describe the swinging motion of the pendulum even though the motion is happening in 2D. The set of coordinates we will use for complicated problems are called *generalized coordinates*. Until we move into the Constrained version of equations of motion, these generalized coordinates will be linearly independent. They need not be physical or easily recognizable.[6]

In general the physical coordinates can be constructed from these generalized coordinates q. For a single particle we could write that transformation as

$$\begin{aligned} x &= h_1(q_1,\ q_2,\ q_3,\ t) \\ y &= h_2(q_1,\ q_2,\ q_3,\ t) \\ z &= h_3(q_1,\ q_2,\ q_3,\ t) \end{aligned} \tag{10}$$

For a system of N particles with l constraints, we need $3N - l$ generalized coordinates. We write the position vector in terms of these generalized coordinates as

$$\boldsymbol{r} = \boldsymbol{r}(\mathbf{q}(t), t) \tag{11}$$

where $\mathbf{q}(t)$ is the set of all the generalized coordinates. It is important to note that this quantity is not a vector[7] but is a set of numbers like a column of a matrix.

[6] When we talk about Hamilton's equations the coordinates we use will involve generalized momenta which is a bit odd the first time you come across it.

[7] It's also not a vector field.

3.2 Generalized forces

We can write the generalized force on the k-th generalized coordinate q_k as

$$Q_k = \sum_l \boldsymbol{F}_l \cdot \frac{\partial \boldsymbol{r}_l}{\partial q_k} + \sum_j \boldsymbol{M}_j \cdot \frac{\partial \boldsymbol{\omega}_j}{\partial \dot{q}_k} \tag{12}$$

where $\boldsymbol{F}_l$ is the l-th force, and $\boldsymbol{r}_l$ is the position vector to the location where the force $\boldsymbol{F}_l$ is applied. The second term covers the external moments, where $\boldsymbol{M}_j$ is the j-th moment and $\boldsymbol{\omega}_j$ is the body's angular velocity about the axis along which the considered moment is applied.

3.3 Lagrange's equation

We can build Lagrange's equation for the k generalized coordinates from

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} + \frac{\partial D}{\partial \dot{q}_k} = Q_k \tag{13}$$

where $L = T - V$ is known as the Lagrangian. The kinetic energy is $T = T(\mathbf{q}, \dot{\mathbf{q}}, t)$, and the potential energy is $V = V(\mathbf{q}, t)$. Rayleigh's dissipation function $D = D(\mathbf{q}, \dot{\mathbf{q}})$ covers simple viscous damping elements (such as dashpots).

3.3.1 Derivation of Lagrange's equation

This proof roughly follows from Meirovitch (2001), with some notation differences. Starting with a system of N particles we can write Newton's 2nd law for the i-th body as

$$\boldsymbol{F}_i + \boldsymbol{f}_i = m_i \ddot{\boldsymbol{r}}_i \tag{14a}$$

where $\boldsymbol{F}_i$ is the resultant external force, $\boldsymbol{f}_i$ is the resultant force from particle interactions, m_i is the *constant* mass, and $\ddot{\boldsymbol{r}}_i$ is the acceleration with respect to an inertial reference. Next, we move the acceleration term to the other side of the equation, and dot the equation with the *virtual displacement* $\delta \boldsymbol{r}_i$. This virtual displacement, also known as the first variation of the displacement, is an *admissible function*[8] representing an arbitrary displacement. For the most part, these variations act like regular differentials in calculus.[9] All this gives the equation

$$(\boldsymbol{F}_i + \boldsymbol{f}_i - m_i \ddot{\boldsymbol{r}}_i) \cdot \delta \boldsymbol{r}_i = 0 \tag{14b}$$

Now we can add together all the N scalar equations, and employing Newton's Third Law all the interacting forces will cancel

$$\sum_{i=1}^{N} (\boldsymbol{F}_i - m_i \ddot{\boldsymbol{r}}_i) \cdot \delta \boldsymbol{r}_i = 0 \tag{14c}$$

This equation is known as d'Alembert's principle. Rearranging things gives

$$\underbrace{\sum_{i=1}^{N} \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i}_{\delta W_{\text{total}}} - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i = 0 \tag{14d}$$

[8] This is a term used in the Calculus of Variations. We can take it to mean that this function is arbitrary but follows necessary boundary conditions.

[9] Lots of hand-waving here, I know. However, we're wading through deep waters and I don't want us to get too lost ☺.

where the term on the left is known as the total virtual work δW_{total}. This virtual work is the sum of both the conservative and non-conservative forces acting on the system

$$\delta W_{\text{total}} = \sum_{i=1}^{N} \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i = \delta W_{\text{con}} + \delta W_{\text{noncon}} \tag{14e}$$

The non-conservative work is

$$\delta W_{\text{noncon}} = \sum_{i=1}^{N} Q_i \delta q_i \tag{14f}$$

where the Q's are defined in (12) and referred to as generalized forces. The conservative work is constructed as

$$\delta W_{\text{con}} = \sum_{i=1}^{N} \boldsymbol{f}_{\text{con}} \cdot \delta \boldsymbol{r}_i \tag{14g}$$

where $\boldsymbol{f}_{\text{con}}$ are the conservative forces.[10] We can write this force as

$$\boldsymbol{f}_{\text{con}} = -\frac{\partial V}{\partial \boldsymbol{r}_i} \tag{14h}$$

where V is the usual mechanical potential energy from elements such as gravity and elastic springs. We want to change the variables from $\boldsymbol{r}$ to $\mathbf{q}$ so we can use the chain rule

$$\frac{\partial V}{\partial \boldsymbol{r}_i} = \frac{\partial V}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \boldsymbol{r}_i} \tag{14i}$$

$$\delta \boldsymbol{r}_i = \frac{\partial \boldsymbol{r}_i}{\partial \mathbf{q}} \delta \mathbf{q} \tag{14j}$$

so that gives us

$$\boldsymbol{f}_{\text{con}} \cdot \delta \boldsymbol{r}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i} \delta \boldsymbol{r}_i = -\left(\frac{\partial V}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \boldsymbol{r}_i}\right)\left(\frac{\partial \boldsymbol{r}_i}{\partial \mathbf{q}} \delta \mathbf{q}\right) = -\frac{\partial V}{\partial \mathbf{q}} \left(\frac{\partial \boldsymbol{r}_i}{\partial \mathbf{q}}\right)^{-1} \frac{\partial \boldsymbol{r}_i}{\partial \mathbf{q}} \delta \mathbf{q} = -\frac{\partial V}{\partial \mathbf{q}} \delta \mathbf{q} = -\delta V \tag{14k}$$

Now the virtual conservative work is just the negative of the variation of the total potential energy.[11]

$$\delta W_{\text{con}} = \sum_{i=1}^{N} \boldsymbol{f}_{\text{con}} \cdot \delta \boldsymbol{r}_i = -\sum_{i=1}^{N} \delta V_i = -\delta V \tag{14l}$$

The next major step is to take d'Alembert's principle, and integrate it over an arbitrary interval of time

$$\int_{t_1}^{t_2} \delta W_{\text{total}} - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \, \mathrm{d}t = 0 \tag{14m}$$

[10]In mathematics a conservative vector field is a vector field that is the gradient of a scalar function. These conservative vector fields have many useful properties such as being path independent and irrotational. We can write down a conservative vector field as $\boldsymbol{f} = \nabla \phi$, where ϕ some scalar function. For conservative forces it is customary to place a negative sign out front so that $\boldsymbol{f}_{\text{con}} = -\nabla V$.

[11]This should not come as a surprise.

Let's take a look at the acceleration term now, and see how that can be restated. We can use integration by parts to reduce the number of derivatives in time.

$$\int_{t_1}^{t_2} \sum_{i=1}^{N} m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \,\mathrm{d}t = \sum_{i=1}^{N} \int_{t_1}^{t_2} m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \,\mathrm{d}t = \sum_{i=1}^{N} \left[m_i \dot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} m_i \dot{\boldsymbol{r}}_i \cdot \delta \dot{\boldsymbol{r}}_i \,\mathrm{d}t \right] \tag{14n}$$

Since the virtual displacements are arbitrary, we can choose them to be $\delta \boldsymbol{r}_i(t_1) = \delta \boldsymbol{r}(t_2) = 0$, which removes the boundary term. As an aside, if we define the kinetic energy of particle i as

$$T_i = \frac{1}{2} m_i \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i. \tag{14o}$$

Then its first variation is

$$\delta T_i = \delta \left(\frac{1}{2} m_i \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i \right) = \frac{1}{2} m_i \delta \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i + \frac{1}{2} m_i \dot{\boldsymbol{r}}_i \cdot \delta \dot{\boldsymbol{r}}_i = m_i \dot{\boldsymbol{r}}_i \cdot \delta \dot{\boldsymbol{r}}_i. \tag{14p}$$

We can see this is the last term in (14n), so putting things back in we can restate (14n)

$$\sum_{i=1}^{N} \left[m_i \dot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} m_i \dot{\boldsymbol{r}}_i \cdot \delta \dot{\boldsymbol{r}}_i \,\mathrm{d}t \right] = \sum_{i=1}^{N} \left[0 - \int_{t_1}^{t_2} \delta T_i \,\mathrm{d}t \right] = - \int_{t_1}^{t_2} \delta T \,\mathrm{d}t \tag{14q}$$

where the variation of the total kinetic energy is δT. Putting what we now know into (14m) gives

$$\int_{t_1}^{t_2} -\delta V + \delta W_{\text{noncon}} + \delta T \,\mathrm{d}t = 0 \tag{14r}$$

This equation is known as Hamilton's principle, and is occasionally used to derive equations of motion in continuous systems.[12] We can now define the useful quantity called the Lagrangian L, and for mechanical systems this is defined as $L \coloneqq T - V$, and we can see the variation of this quantity in Hamilton's principle.

$$\int_{t_1}^{t_2} \delta L + \sum_{i=1}^{N} Q_i \delta q_i \,\mathrm{d}t = 0 \tag{14s}$$

In general we wish to construct L in terms of a coordinates $\mathbf{q} = (q_1, q_2, \ldots, q_n)$, and their derivatives $\dot{\mathbf{q}}$. For the moment, let's only consider the minimum set of generalized coordinates so they must be independent.[13] We can usually think of these as the variables that are used to parameterize or construct the position vectors. In the most general case we would have as a function of $\mathbf{q}$ and $\dot{\mathbf{q}}$.

$$L \coloneqq T(\mathbf{q}, \dot{\mathbf{q}}, t) - V(\mathbf{q}, t) = L(\mathbf{q}, \dot{\mathbf{q}}, t) \tag{14t}$$

Expanding the variation of L gives

$$\delta L = \sum_{i=1}^{N} \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] \tag{14u}$$

[12] Naturally, it can be used in discrete systems too, but it is usually more work than Lagrange's equation. It can be handy, however, in strange cases.

[13] No redundant coordinates with additional constraints.

Placing this back into (14s) gives

$$\sum_{i=1}^{N}\left[\int_{t_1}^{t_2}\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i+Q_i\delta q_i\,\mathrm{d}t\right]=0 \tag{14v}$$

Now we just need to get $\delta\dot{q}_i$ in terms of δq_i and we're home free. To do this, we will again integrate by parts.

$$\int_{t_1}^{t_2}\frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\,\mathrm{d}t=\left.\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right|_{t_1}^{t_2}-\int_{t_1}^{t_2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i\,\mathrm{d}t \tag{14w}$$

Just like before, we can arbitrarily set $\delta q_i(t_1)=\delta q_i(t_2)=0$ which removes the boundary terms. Placing this result back into (14v) gives

$$\int_{t_1}^{t_2}\sum_{i=1}^{N}\left[\frac{\partial L}{\partial q_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)+Q_i\right]\delta q_i\,\mathrm{d}t=0 \tag{14x}$$

The only way with arbitrary δq_i that all these integrals are identically equal to zero is if each coefficient of δq_i is zero. This gives a set of N equations, know generally as the Euler-Lagrange equations.

$$\frac{\partial L}{\partial q_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)+Q_i=0,\quad i=1,\,2,\,\dots\,,N \tag{14y}$$

In engineering we often rearrange the terms to place the generalized forces on the right-hand side of the equation. Also, linear viscous damping is commonly treated with Rayleigh's dissipation function. The end result, as we usually write it, is given in (13).

3.3.2 Example: pendulum

Let's consider the usual simple pendulum, shown in Figure 1. From the hinge the position vector of the mass is

$$\begin{aligned}\boldsymbol{r}&=l\left(\sin\theta\hat{\boldsymbol{i}}-\cos\theta\hat{\boldsymbol{j}}\right) &\text{(15a)}\\ &=l\hat{\boldsymbol{e}}_r &\text{(15b)}\end{aligned}$$

We can use either frame, as long as we are careful when computing derivatives. I'll pick my single generalized coordinate to be $q=\theta$. Computing the velocity

$$\begin{aligned}\dot{\boldsymbol{r}}&=l\dot{\theta}\left(\cos\theta\hat{\boldsymbol{i}}+\sin\theta\hat{\boldsymbol{j}}\right) &\text{(15c)}\\ &=l\dot{\theta}\hat{\boldsymbol{k}}\times\hat{\boldsymbol{e}}_r=l\dot{\theta}\hat{\boldsymbol{e}}_\theta &\text{(15d)}\end{aligned}$$

Now I'll compute the kinetic energy

$$\begin{aligned}T&=\frac{1}{2}m\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} &\text{(15e)}\\ &=\frac{1}{2}ml^2\dot{\theta}^2\left(\cos\theta\hat{\boldsymbol{i}}+\sin\theta\hat{\boldsymbol{j}}\right)\cdot\left(\cos\theta\hat{\boldsymbol{i}}+\sin\theta\hat{\boldsymbol{j}}\right)=\frac{1}{2}ml^2\dot{\theta}^2\left(\cos^2\theta+\sin^2\theta\right)=\frac{1}{2}ml^2\dot{\theta}^2 &\text{(15f)}\\ &=\frac{1}{2}m\left(l\dot{\theta}\hat{\boldsymbol{e}}_\theta\right)\cdot\left(l\dot{\theta}\hat{\boldsymbol{e}}_\theta\right)=\frac{1}{2}ml^2\dot{\theta}^2 &\text{(15g)}\end{aligned}$$

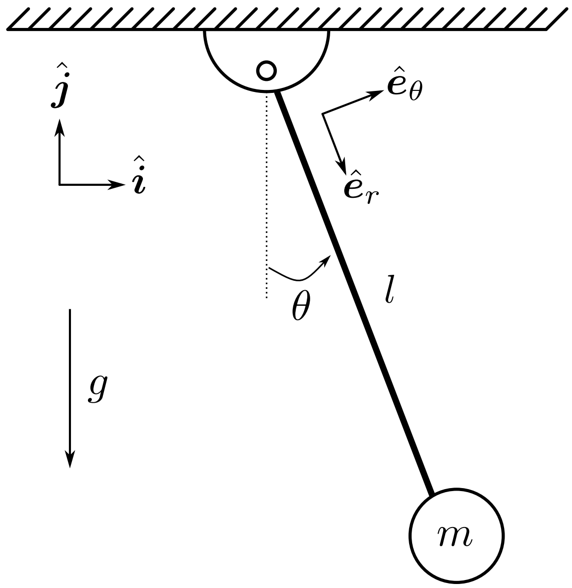

Figure 1: A simple planar pendulum.

Next we can construct the potential energy V, and we can see that the only potential comes from gravity. I'll define the datum to be at the hinge, which gives

$$V = mg\boldsymbol{r} \cdot \hat{\boldsymbol{j}} = -mgl\cos\theta \tag{15h}$$

Now we build the Lagrangian

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta \tag{15i}$$

and note that there are no external forces acting on the system so $Q = 0$. Finally we use Lagrange's equation to arrive at the equation of motion.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0 \tag{15j}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(ml^2\dot{\theta}\right) - (-mgl\sin\theta) = 0 \tag{15k}$$

$$ml^2\ddot{\theta} + mgl\sin\theta = 0 \tag{15l}$$

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0 \tag{15m}$$

3.4 Hamilton's equation

The Legendre transform of the Lagrangian $L(q, \dot{q})$ is known as the Hamiltonian $\mathcal{H}(q, p)$. Through this canonical transform, we can rewrite Lagrange's equations as a set of first-order differential equations known as Hamilton's equations. However, the coordinates we use to solve these problems are different. Instead of $(q, \dot{q})$ we have (q, p), where the generalized momenta[14] p is defined as

$$p_k = \frac{\partial L}{\partial \dot{q}_k}. \tag{16}$$

[14] Sometimes called the conjugate momenta

This formulation has some very nice properties for both analysis (like finding conserved states), as well as numerics (having a system in explicit first-order form for solvers).

For mechanical systems, the Lagrangian has a form that when transformed usually becomes the total energy. This is so common many take $\mathcal{H} = T + V$ as a definition. This works fine most of the time, but it is not universally true. To build Hamilton's equation from our typical mechanical systems:

1. We first build T, V, L, and Q_k as we did for Lagrange's equation, using generalized coordinates $(q, \dot{q})$.
2. Then compute (16) for each coordinate, and solve them simultaneously for $\dot{q}_k$.
3. Build the Hamiltonian

$$\mathcal{H} = T + V \tag{17}$$

and substitute $\dot{q}_k$ in terms of the previous step to get $\mathcal{H}(q, p)$.[15]

4. Finally, we build Hamilton's equations

$$\dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k} + Q_k \tag{18a}$$

$$\dot{q}_k = \frac{\partial \mathcal{H}}{\partial p_k} \tag{18b}$$

We can think of Q_k as also containing all the non-conservative forces as well, such as those due to damping. If we had $D(q, \dot{q})$ from Lagrange's equation, we can solve for the force from this term

$$Q_{k,\text{ damping}} = -\frac{\partial D}{\partial \dot{q}_k} \tag{19}$$

and then transform the results from $\dot{q}$ to p as we did for the Hamiltonian.

3.4.1 Derivation of Hamilton's equation

If we take the most general case, then the Lagrangian is a function

$$L = L(\mathbf{q}, \dot{\mathbf{q}}, t) \tag{20a}$$

and the total differential of L is

$$\mathrm{d}L = \sum_{i=1}^{N} \left(\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i + \frac{\partial L}{\partial \dot{q}_i} \,\mathrm{d}\dot{q}_i \right) + \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20b}$$

We can use the definition of (16) to have

$$\mathrm{d}L = \sum_{i=1}^{N} \left(\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i + p_i \,\mathrm{d}\dot{q}_i \right) + \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20c}$$

As an aside we can see that

$$\mathrm{d}\left(p_i \dot{q}_i\right) = \mathrm{d}p_i \dot{q}_i + p_i \,\mathrm{d}\dot{q}_i \quad \rightarrow \quad p_i \,\mathrm{d}\dot{q}_i = \mathrm{d}\left(p_i \dot{q}_i\right) - \mathrm{d}p_i \dot{q}_i \tag{20d}$$

[15] As brought up later, the Hamiltonian is technically not always the total energy $T + V$. It is actually the Legendre transform of the Lagrangian (20g) and under certain conditions, that are very common in mechanical problems, this becomes the total energy.

so we can substitute that back into $\mathrm{d}L$ to remove the $\mathrm{d}\dot{q}_i$ terms.

$$\mathrm{d}L = \sum_{i=1}^{N} \left(\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i + \mathrm{d}\left(p_i \dot{q}_i\right) - \dot{q}_i \,\mathrm{d}p_i \right) + \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20e}$$

Rearranging terms gives us

$$\sum_{i=1}^{N} \mathrm{d}\left(p_i \dot{q}_i\right) - \mathrm{d}L = -\sum_{i=1}^{N} \left(\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i - \dot{q}_i \,\mathrm{d}p_i \right) - \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20f}$$

Now, let us define the Hamiltonian, as the Legendre Transform of the Lagrangian

$$\mathcal{H} \coloneqq \sum_{i=1}^{N} p_i \dot{q}_i - L \tag{20g}$$

We can see that the left hand side of the previous equation is the differential of the Hamiltonian

$$\mathrm{d}\mathcal{H} = \mathrm{d}\left(\sum_{i=1}^{N} p_i \dot{q}_i - L \right) = \sum_{i=1}^{N} \left(-\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i + \dot{q}_i \,\mathrm{d}p_i \right) - \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20h}$$

In general, the Hamiltonian is a function $\mathcal{H}(q, p, t)$, so if we compute the total differential of this function we have the following by definition

$$\mathrm{d}\mathcal{H} = \mathrm{d}\mathcal{H}(q, p, t) = \sum_{i=1}^{N} \left(\frac{\partial \mathcal{H}}{\partial q_i} \,\mathrm{d}q_i + \frac{\partial \mathcal{H}}{\partial p_i} \,\mathrm{d}p_i \right) + \frac{\partial \mathcal{H}}{\partial t} \,\mathrm{d}t. \tag{20i}$$

We can observe that the $\mathrm{d}\mathcal{H}$ in both (20h) and (20i) must be the same, so we can equate these equations and collect terms.

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{N} \left(\frac{\partial \mathcal{H}}{\partial q_i} \,\mathrm{d}q_i + \frac{\partial \mathcal{H}}{\partial p_i} \,\mathrm{d}p_i \right) + \frac{\partial \mathcal{H}}{\partial t} \,\mathrm{d}t = \sum_{i=1}^{N} \left(-\frac{\partial L}{\partial q_i} \,\mathrm{d}q_i + \dot{q}_i \,\mathrm{d}p_i \right) - \frac{\partial L}{\partial t} \,\mathrm{d}t \tag{20j}$$

$$\sum_{i=1}^{N} \left(\frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial L}{\partial q_i} \right) \mathrm{d}q_i + \sum_{i=1}^{N} \left(\frac{\partial \mathcal{H}}{\partial p_i} - \dot{q}_i \right) \mathrm{d}p_i + \left(\frac{\partial \mathcal{H}}{\partial t} + \frac{\partial L}{\partial t} \right) \mathrm{d}t = 0 \tag{20k}$$

Using a similar arbitrariness argument to the derivation of Lagrange's equation, we can say that the only way to ensure the previous equation is always identically zero is if each coefficient of $\mathrm{d}q_i$, $\mathrm{d}p_i$, and $\mathrm{d}t$ must be zero. This gives $2N+1$ equations

$$\frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial L}{\partial q_i} = 0, \quad i \in \{1, 2, \ldots, N\} \tag{20l}$$

$$\frac{\partial \mathcal{H}}{\partial p_i} - \dot{q}_i = 0, \quad i \in \{1, 2, \ldots, N\} \tag{20m}$$

$$\frac{\partial \mathcal{H}}{\partial t} + \frac{\partial L}{\partial t} = 0 \tag{20n}$$

Taking (20l) we can replace $\frac{\partial L}{\partial q_i}$ by using Lagrange's equation (14y) for coordinate q_i

$$\frac{\partial L}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - Q_i \tag{20o}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t} \left(p_i\right) - Q_i = \dot{p}_i - Q_i \tag{20p}$$

Now we can put this back into (20l)

$$\frac{\partial \mathcal{H}}{\partial q_i} + \dot{p}_i - Q_i = 0 \tag{20q}$$

Rearranging this and (20m) we arrive at the general form of Hamilton's equations

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} + Q_i \tag{20r}$$

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{20s}$$

$$i \in \{1, 2, \ldots, N\}$$

along with the auxilary equation that relates the time rate of change of the Hamiltonian and the Lagrangian

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial L}{\partial t}. \tag{20t}$$

3.4.2 The Hamiltonian

(Zia, Redish, and McKay, 2009) (Goldstein, Poole, and Safko, 2001, Chap. 2.7)

It's very common for our mechanical systems that the Hamiltonian is the total energy $\mathcal{H} = T + V$. However, this is not universally true, particularly in some modern physics problems. The definition of (20g) is universally true. Let's take a look at the condition for when the Hamiltonian is the total energy.

complete the derivation about total energy

3.4.3 Example: pendulum

Again, let's consider the usual simple pendulum, shown in Figure 1. From the Lagrange's equation example, we known the Lagrangian to be

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta. \tag{21a}$$

where the generalized coordinate is $q = \theta$. So we can rewrite the Lagrangian as

$$L = \frac{1}{2}ml^2\dot{q}^2 + mgl\cos q. \tag{21b}$$

The next step is to compute the generalized momentum p, and invert it for $\dot{q}$.

$$p = \frac{\partial L}{\partial \dot{q}} = ml^2\dot{q} \quad \rightarrow \quad \dot{q} = \frac{p}{ml^2} \tag{21c}$$

Now we need to build the Hamiltonian, and I'll use the general form of the Legendre transform this time to show how it works.

$$\mathcal{H} = \sum_{i=1}^{N} p_i\dot{q}_i - L = p\dot{q} - L = p\dot{q} - \frac{1}{2}ml^2\dot{q}^2 - mgl\cos q \tag{21d}$$

$$= p\left(\frac{p}{ml^2}\right) - \frac{1}{2}ml^2\left(\frac{p}{ml^2}\right)^2 - mgl\cos q \tag{21e}$$

$$= \left(\frac{p^2}{ml^2}\right) - \frac{1}{2}\left(\frac{p^2}{ml^2}\right) - mgl\cos q \tag{21f}$$

Simplifying the first two terms gives the Hamiltonian that'll we'll use for the next step. Note that this is now in terms of generalized coordinate q and generalized momentum p.

$$\mathcal{H} = \frac{1}{2}\frac{p^2}{ml^2} - mgl\cos q \tag{21g}$$

Also we can see that the final result is $\mathcal{H} = T + V$ with $\dot{q} \to p^2/(ml^2)$. Since we now have the Hamiltonian, we can directly use (18) to build the 2 first order equations.

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial q} + \cancelto{0}{Q} = -mgl \sin q \tag{21h}$$

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{ml^2} \tag{21i}$$

3.5 Constrained Lagrange's equation

Write about constrained version of Lagrange's equations

3.5.1 Example: Block on an incline

3.5.2 Example: Pendulum

Include example of pendulum

3.6 Udwadia–Kalaba equation

3.6.1 Overview

This formulation of equations of motion is useful if the system is complicated, and constrained. First, consider the uncoupled body, and define a set of n coordinates $\mathbf{q}$ that define the position of the system. The full Udwadia-Kalaba equation of the constrained motion is

$$\mathbf{M}\ddot{\mathbf{q}} = \mathbf{Q} + \mathbf{M}^{1/2}\left(\mathbf{A}\mathbf{M}^{-1/2}\right)^{+}\left(\mathbf{b} - \mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\right) \tag{24}$$

where

- $\mathbf{M} \in \mathbb{R}^{n \times n}$ is the mass matrix of the unconstrained system. We can usually build this directly, or use the unconstrained kinetic energy $M_{ij} = \frac{\partial^2 T}{\partial \dot{q}_i \partial \dot{q}_j}$.
- $\mathbf{Q} \in \mathbb{R}^n$ is the column of external forces. If we have some quantities known in terms of the potential energy we could build

$$\mathbf{Q} = -\frac{\partial V}{\partial \mathbf{q}} - \frac{\partial D}{\partial \dot{\mathbf{q}}} + \mathbf{Q}_{\text{other}} \tag{25}$$

- $\mathbf{A}$ and $\mathbf{b}$ are related to the constraints. The constraints must be of the form

$$\mathbf{A}(\mathbf{q}, \dot{\mathbf{q}}, t)\ddot{\mathbf{q}} = \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}, t) \tag{26}$$

  It is important to remember that this is constraining the accelerations, not the positions or velocities. Which means if we choose a poor time-marching scheme, we will see constraint drift (the actual constraints we care about will not longer be satisfied).

- $()^+$ means the *Moore–Penrose inverse*, also called the pseudoinverse. Since the matrix $\mathbf{A} \in \mathbb{R}^{l,n}$, where the number of constraints $l < n$, we could explicitly write the pseudoinverse as

$$\left(\mathbf{A}\mathbf{M}^{-1/2}\right)^{+} = \mathbf{P}^{+} = \mathbf{P}^T\left(\mathbf{P}\mathbf{P}^T\right)^{-1}$$

  However, we should usually let this be computed numerically. Both Matlab[16] and Julia[17] have standard implementations of the pseudoinverse that use the singular value decomposition, which is very numerically robust.

[16] Matlab's built-in Moore-Penrose inverse is `pinv`

[17] Julia's built-in Moose-Penrose inverse is `pinv`, but the backslash operator also will automatically compute the pseudo-inverse.

- $\mathbf{M}^{1/2}$ means the *square root of a matrix*. This means we would need to compute a matrix such that

  $$\mathbf{M}^{1/2}\mathbf{M}^{1/2} = \mathbf{M}.$$

  Several numerical techniques have been designed to do this, either by iteration, the singular value decomposition, or the Schur decomposition. Most current implementations use a Schur decomposition, such as Matlab's[18] implementation of Deadman, Higham, and Ralha (2013), while Julia[19] uses Björck and Hammarling (1983).

  Typically, $\mathbf{M}$ is symmetric and non-negative definite.[20] If we compute the singular value decomposition of $\mathbf{M} = \mathbf{USV}^T$, we would have $\mathbf{U} = \mathbf{V}$ because $\mathbf{M}$ is non-negative definite. We can build the matrix

  $$\mathbf{M}^{1/2} = \mathbf{VS}^{1/2}\mathbf{V}^T \tag{27}$$

  where $\mathbf{S}^{1/2} = \mathrm{diag}(\sigma_1^{1/2}, \ldots, \sigma_n^{1/2})$. It's direct to see that $\mathbf{M}^{1/2}\mathbf{M}^{1/2} = \mathbf{VSV}^T = \mathbf{M}$, and that $\mathbf{M}^{1/2}$ is symmetric.

3.6.2 Building the constraint matrices

In most mechanical systems, the constraints will likely first be formulated as a combination of holonomic and non-holonomic constraints. A geometric, or *holonomic*, constraint looks like

$$\phi(q, t) = 0 \tag{28}$$

They come from things like requiring the length of a pendulum to be fixed.

The other type of constraint is one that is velocity-dependent, also known as a *non-holonomic* constraint. These have the form

$$\psi(q, \dot{q}, t) = 0 \tag{29}$$

They come from things like requiring the velocity of a slider to point in a particular direction.

A particular system may have some combination of multiple types of constraints. Each constraint becomes a row in $\mathbf{A}$ and has an associated entry in $\mathbf{b}$. To compute these entries, we can differentiate ϕ twice and ψ once, with respect to time. Using the chain rule

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\partial \phi}{\partial t} + \sum_i \frac{\partial \phi}{\partial q_i}\dot{q}_i = 0$$

and differentiating again gives

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} = \frac{\partial^2 \phi}{\partial t^2} + \sum_i \frac{\partial \phi}{\partial q_i}\ddot{q}_i + \sum_j \sum_i \dot{q}_j \frac{\partial^2 \phi}{\partial q_i \partial q_j}\dot{q}_i = 0$$

We can rearrange this expression to have the form of a single row of (26).

$$\begin{bmatrix} \frac{\partial \phi}{\partial q_1} & \frac{\partial \phi}{\partial q_2} & \cdots & \frac{\partial \phi}{\partial q_n} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_n \end{bmatrix} = -\frac{\partial^2 \phi}{\partial t^2} - \begin{bmatrix} \dot{q}_1 & \dot{q}_2 & \cdots & \dot{q}_n \end{bmatrix} \mathbf{H} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_n \end{bmatrix}$$

[18] Matlab documentation on `sqrtm`

[19] Julia documentation on `sqrt` in the `LinearAlgebra` standard library.

[20] All the eigenvalues are real, and greater than or equal to zero.

where $\mathbf{H} = \frac{\partial^2 \phi}{\partial \mathbf{q} \partial \mathbf{q}}$ is the Hessian of ϕ. We can write this a little more compactly, as

$$\left(\frac{\partial \phi}{\partial \mathbf{q}}\right)^T \ddot{\mathbf{q}} = -\frac{\partial^2 \phi}{\partial t^2} - \dot{\mathbf{q}}^T \frac{\partial^2 \phi}{\partial \mathbf{q} \partial \mathbf{q}} \dot{\mathbf{q}} \tag{30}$$

This means that if ϕ was constraint j of the system then the j-th row of $\mathbf{A}$ is $\left(\frac{\partial \phi}{\partial \mathbf{q}}\right)^T$, and the j-th entry in $\mathbf{b}$ is the right hand side of (30).

For the non-holonomic case, we only need to differentiate a single time.

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = 0 = \frac{\partial \psi}{\partial t} + \sum_i \frac{\partial \psi}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial \psi}{\partial \dot{q}_i} \ddot{q}_i$$

Rearranging we get

$$\left(\frac{\partial \psi}{\partial \dot{\mathbf{q}}}\right)^T \ddot{\mathbf{q}} = -\frac{\partial \psi}{\partial t} - \left(\frac{\partial \psi}{\partial \mathbf{q}}\right)^T \dot{\mathbf{q}} \tag{31}$$

which is the same form as (26)

3.6.3 Proof that Udwadia–Kalaba satisfies constraints

We can show that the Udwadia–Kalaba equation (24)

$$\mathbf{M}\ddot{\mathbf{q}} = \mathbf{Q} + \mathbf{M}^{1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \left(\mathbf{b} - \mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\right) \tag{32a}$$

satisfies the constraint equations (26), $\mathbf{A}\ddot{\mathbf{q}} = \mathbf{b}$. I'll take the equation of motion and multiply through by $\mathbf{M}^{-1}$, and define $\mathbf{a} = \mathbf{M}^{-1}\mathbf{Q}$

$$\begin{aligned} \ddot{\mathbf{q}} &= \mathbf{M}^{-1}\mathbf{Q} + \mathbf{M}^{-1}\mathbf{M}^{1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \left(\mathbf{b} - \mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\right) && \text{(32b)} \\ &= \mathbf{a} + \mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ (\mathbf{b} - \mathbf{A}\mathbf{a}) && \text{(32c)} \end{aligned}$$

Now, we can multiply through by $\mathbf{A}$ and expand

$$\begin{aligned} \mathbf{A}\ddot{\mathbf{q}} &= \mathbf{A}\mathbf{a} + \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ (\mathbf{b} - \mathbf{A}\mathbf{a}) && \text{(32d)} \\ &= \mathbf{A}\mathbf{a} + \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \mathbf{b} - \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \mathbf{A}\mathbf{a} && \text{(32e)} \\ &= \left(\mathbf{I} - \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+\right) \mathbf{A}\mathbf{a} + \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \mathbf{b} && \text{(32f)} \\ &= \left(\mathbf{I} - \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+\right) \mathbf{A}\mathbf{M}^{-1/2}\mathbf{M}^{1/2}\mathbf{a} + \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+ \mathbf{b} && \text{(32g)} \end{aligned}$$

where we multiply the first term by $\mathbf{I} = \mathbf{M}^{-1/2}\mathbf{M}^{1/2}$ in the last step. Next we rely on one of the identities of the pseudo-inverse

$$\mathbf{P}\mathbf{P}^+\mathbf{P} = \mathbf{P}. \tag{32h}$$

If we let $\mathbf{P} = \mathbf{A}\mathbf{M}^{-1/2}$, then the first term becomes

$$\begin{aligned} \left(\mathbf{I} - \mathbf{A}\mathbf{M}^{-1/2} \left(\mathbf{A}\mathbf{M}^{-1/2}\right)^+\right) \mathbf{A}\mathbf{M}^{-1/2}\mathbf{M}^{1/2}\mathbf{a} &= \left(\mathbf{I} - \mathbf{P}\mathbf{P}^+\right) \mathbf{P}\mathbf{M}^{1/2}\mathbf{a} && \text{(32i)} \\ &= \left(\mathbf{P} - \mathbf{P}\mathbf{P}^+\mathbf{P}\right) \mathbf{M}^{1/2}\mathbf{a} && \text{(32j)} \\ &= (\mathbf{P} - \mathbf{P})\, \mathbf{M}^{1/2}\mathbf{a} = \mathbf{0} && \text{(32k)} \end{aligned}$$

This simplifies the previous result quite a bit, now we can rewrite (32g) as

$$\mathbf{A}\ddot{\mathbf{q}} = \mathbf{A}\mathbf{M}^{-1/2}\left(\mathbf{A}\mathbf{M}^{-1/2}\right)^{+}\mathbf{b} \tag{32l}$$

We can continue to work on this. Since the matrix $\mathbf{P} = \left(\mathbf{A}\mathbf{M}^{-1/2}\right)$ has size $l \times n$ (and $l < n$), it has linearly independent rows and therefore is said to have a right-inverse $\mathbf{P}\mathbf{P}^{+} = \mathbf{I}$. This simplifies the right hand side to $\mathbf{b}$.

$$\mathbf{A}\ddot{\mathbf{q}} = \underbrace{\left(\mathbf{A}\mathbf{M}^{-1/2}\right)\left(\mathbf{A}\mathbf{M}^{-1/2}\right)^{+}}_{=\mathbf{I}}\mathbf{b} \tag{32m}$$

$$\mathbf{A}\ddot{\mathbf{q}} = \mathbf{b} \tag{32n}$$

This completes the derivation that the Udwadia–Kalaba equation satisfies all the constraints.

3.6.4 Example: Pendulum

Let's consider a good old planar pendulum, as shown in Figure 1.

We can use the coordinates $\mathbf{q} = [x, y]^T$ of the unconstrained coordinates to describe the motion. The constraint is a geometric one, namely the length of the pendulum cannot change. This means we have a single $\phi(\mathbf{q})$ for the distance between the hinge and the center of mass.

$$\phi = x^2 + y^2 - L^2 = q_1^2 + q_2^2 - L^2 \tag{33a}$$

Using (30) we get

$$\underbrace{\begin{bmatrix} 2q_1 & 2q_2 \end{bmatrix}}_{\mathbf{A}}\ddot{\mathbf{q}} = \underbrace{0 - \begin{bmatrix} \dot{q}_1 & \dot{q}_2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix}}_{b} \tag{33b}$$

This gives

$$\mathbf{A} = \begin{bmatrix} 2q_1 & 2q_2 \end{bmatrix} \tag{33c}$$

$$b = -2\dot{q}_1^2 - 2\dot{q}_2^2 \tag{33d}$$

Since there is only a single constraint, then $\mathbf{b} \to b$ (it is a scalar).

The only external force acting on the unconstrained particle is gravity. We could pose this in terms of the potential energy $V = mgy$, or just directly see the force is

$$\mathbf{Q} = \begin{bmatrix} 0 \\ -mg \end{bmatrix} \tag{33e}$$

The mass matrix is directly computed from the kinetic energy of the uncontrained particle

$$T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 \tag{33f}$$

therefore

$$\mathbf{M} = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \tag{33g}$$

We now have built all the parts of (24), and can directly implement the equations numerically. To illustrate that these equations are the same as the equation derived from the constrained form of Lagrange's equation, we shall continue analytically. However, this is not commonly done in larger systems. Analytically, I'll begin with the square root, which is simple since $\mathbf{M}$ is diagonal

$$\mathbf{M}^{1/2} = \begin{bmatrix} \sqrt{m} & 0 \\ 0 & \sqrt{m} \end{bmatrix}$$

then

$$\mathbf{P} = \mathbf{A}\mathbf{M}^{-1/2} = \begin{bmatrix} \frac{2}{\sqrt{m}}q_1 & \frac{2}{\sqrt{m}}q_2 \end{bmatrix}$$

$$\mathbf{P}^{+} = \mathbf{P}^T \left(\mathbf{P}\mathbf{P}^T\right)^{-1} = \frac{\sqrt{m}}{2\left(q_1^2 + q_2^2\right)} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$

Continuing the parts on the right

$$\left(b - \mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\right) = -2\dot{q}_1^2 - 2\dot{q}_2^2 + 2gq_2$$

Putting this all back into the right-hand side of (24)

$$\mathbf{Q} + \mathbf{M}^{1/2}\left(\mathbf{A}\mathbf{M}^{-1/2}\right)^{+}\left(b - \mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\right)$$
$$= \begin{bmatrix} 0 \\ -mg \end{bmatrix} + \begin{bmatrix} \sqrt{m} & 0 \\ 0 & \sqrt{m} \end{bmatrix} \left(\frac{\sqrt{m}}{2\left(q_1^2 + q_2^2\right)} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}\right)\left(-2\dot{q}_1^2 - 2\dot{q}_2^2 + 2gq_2\right)$$

Lastly, solving for $\ddot{\mathbf{q}}$ explicitly and simplifying gives

$$\begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -g \end{bmatrix} + \frac{gq_2 - \dot{q}_1^2 - \dot{q}_2^2}{q_1^2 + q_2^2} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \tag{33h}$$

References

Baruh, H. (1999). *Analytical Dynamics*. McGraw-Hill.

— (2014). *Applied Dynamics*. CRC Press.

Björck, Å. and S. Hammarling (1983). “A Schur method for the square root of a matrix”. *Linear Algebra and its Applications*, **52-53**: pp. 127–140. DOI: 10.1016/0024-3795(83)80010-x.

Deadman, E., N. J. Higham, and R. Ralha (2013). “Blocked Schur Algorithms for Computing the Matrix Square Root”. In: *Lecture Notes in Computer Science*. Vol. 7782: *Applied Parallel and Scientific Computing, PARA 2012*. Ed. by P. Manninen and P. Öster. Springer, pp. 171–182. DOI: 10.1007/978-3-642-36803-5_12.

Goldstein, H., C. P. Poole, and J. L. Safko (2001). *Classical Mechanics*. 3rd ed. Pearson.

Greenwood, D. T. (1977). *Classical Dynamics*. Dover.

— (2003). *Advanced Dynamics*. Cambridge University Press.

Kane, T. R. and D. A. Levinson (1985). *Dynamics, Theory and Applications*. McGraw-Hill.

Marques, F., A. P. Souto, and P. Flores (2016). “On the constraints violation in forward dynamics of multibody systems”. *Multibody System Dynamics*, **39**:4, pp. 385–419. DOI: 10.1007/s11044-016-9530-y.

Meirovitch, L. (2001). *Fundamentals of Vibrations*. McGraw-Hill.

Roithmayr, C. M. and D. H. Hodges (2016). *Dynamics: Theory and Application of Kane’s Method*. Cambridge University Press.

Schaub, H. and J. L. Junkins (2014). *Analytical Mechanics of Space Systems*. 3rd ed. American Institute of Aeronautics and Astronautics, Inc. DOI: 10.2514/4.102400.

Udwadia, F. E. and R. E. Kalaba (1992). “A new perspective on constrained motion”. *Proceedings: Mathematical and Physical Sciences*, pp. 407–410.

— (2007). *Analytical Dynamics: A New Approach*. Cambridge University Press.

Udwadia, F. E. and A. D. Schutte (2010). “Equations of motion for general constrained systems in Lagrangian mechanics”. *Acta Mechanica*, **213**:1-2, pp. 111–129. DOI: 10.1007/s00707-009-0272-2.

Zia, R. K. P., E. F. Redish, and S. R. McKay (2009). “Making sense of the Legendre transform”. *American Journal of Physics*, **77**:7, pp. 614–622. DOI: 10.1119/1.3119512.