Lecture Notes on Classical Mechanics for Physics 106ab – Errata

Sunil Golwala

Revision Date: January 15, 2007

Revision Date: January 15, 2007

This document lists errors in the September 25, 2006, version of the lecture notes. Corrections are listed by section. Many thanks to students who have informed me of errors and typos.

Changes made between Dec 29 and Jan 15:

- Corrections added for Chapter 5 on Rotating Systems. Perhaps incomplete; more corrections may be added later.

Contents

Chapter 1

Elementary Mechanics

1.1 Newtonian Mechanics

1.1.3 Energy and Work

Potential Energy, Conservation of Energy, and Conservative Forces

- In the discussion of forces for which the work W_{12} is path-independent, we should replace

> Furthermore, consider forces for which W_{12} is *path-independent*, *i.e.*, depends only on $\vec{r}_1$ and $\vec{r}_2$. (Do there exist position-dependent forces for which this is not true? Hard to think of any.)

by

> Furthermore, consider forces for which W_{12} is *path-independent*, *i.e.*, depends only on $\vec{r}_1$ and $\vec{r}_2$. Another way of saying this is that the work done around a closed path vanishes: pick any two points 1 and 2, calculate the work done in going from 1 to 2 and from 2 to 1. The latter will be the negative of the former if the work done is path-independent. By Stokes' Theorem (see Appendix A), we then see that path-independence of work is equivalent to requiring that $\vec{\nabla} \times \vec{F} = 0$ everywhere. (Do there exist position-dependent forces for which this is not true? Hard to think of any physically realized ones, but one can certainly construct force functions with nonzero curl.)

Calculating Motion from the Potential Energy

- In Example 1.9, after we solve for the equilibrium position, we say

> There is a stable equilibrium if $m_1 > m_2/2$; otherwise the square root becomes zero or imaginary.

This is incomplete; please replace with:

> There is an equilibrium if $m_1 > m_2/2$ (so that the square root is neither zero nor imaginary) and if b and d are such that the resulting value of $z_1 < 0$: m_1 is not allowed to go above point B.

1.3 Dynamics of Systems of Particles

1.3.1 Newtonian Mechanical Concepts for Systems of Particles

Angular Momentum, Conservation of Angular Momentum, External and Internal Torques

- In the calculation of the total angular momentum $\vec{L}$, we have the line

$$= \sum_a m_a \left[\left(\vec{R} \times \dot{\vec{R}} \right) + \left(\vec{R} \times \dot{\vec{s}}_a \right) + \left(\vec{s}_a \times \dot{\vec{R}} \right) + \left(\vec{s}_a \times \dot{\vec{s}}_a \right) \right]$$

  which should of course be

$$= \sum_a m_a \left[\left(\vec{R} \times \dot{\vec{R}} \right) + \left(\vec{R} \times \dot{\vec{s}}_a \right) + \left(\vec{s}_a \times \dot{\vec{R}} \right) + \left(\vec{s}_a \times \dot{\vec{s}}_a \right) \right]$$

1.3.2 The Virial Theorem

We add a new section on the Virial Theorem, between the original Sections 1.3.1 (*Newtonian Mechanical Concepts for Systems of Particles*) and 1.3.2 (*Collisions of Particles*)

Here we prove the Virial Theorem, which relates the time-averaged kinetic energy for a bounded system to a quantity called the **virial**, which is just a time-averaged dot product of the force and position of the various particles in the system. In its basic form, the virial theorem does not have a clear intuitive interpretation, though it is certainly useful. When one considers the specific case of conservative forces that depend on particle radius, the virial becomes simply related to the potential energy of the system. Thus, we obtain a time-averaged relation between kinetic and potential energy. This is an incredibly powerful statement because it doesn't require specific knowledge of the particle orbits.

Generic Version

Consider an ensemble of particles, whose positions $\vec{r}_a$ and momenta $\vec{p}_a$ are bounded, meaning that there are upper limits on both. This means that the particles are both confined to a particular region of space and also that they never approach a force center that might impart to them infinite momentum. Define the quantity

$$S = \sum_a \vec{p}_a \cdot \vec{r}_a$$

Calculate the time-averaged rate of change of S:

$$\left\langle \frac{dS}{dt} \right\rangle = \frac{1}{\tau} \int_0^\tau \frac{dS}{dt}\, dt$$

The integrand is a total derivative, so the integral is done trivially:

$$\left\langle \frac{dS}{dt} \right\rangle = \frac{S(\tau) - S(0)}{\tau}$$

Since we assumed $\vec{r}_a$ and $\vec{p}_a$ are bounded, it also holds that S is bounded. Thus, by letting $\tau \to \infty$ – *i.e.*, by taking the average over an arbitrarily long

time – we can make $\left\langle \frac{dS}{dt} \right\rangle \to 0$. Let us explicitly calculate $\left\langle \frac{dS}{dt} \right\rangle$ and use the fact that it vanishes:

$$0 = \left\langle \frac{dS}{dt} \right\rangle = \left\langle \sum_a \left[\vec{p}_a \cdot \dot{\vec{r}}_a + \dot{\vec{p}}_a \cdot \vec{r}_a \right] \right\rangle = \left\langle \sum_a \vec{p}_a \cdot \dot{\vec{r}}_a \right\rangle + \left\langle \sum_a \dot{\vec{p}}_a \cdot \vec{r}_a \right\rangle$$

The two terms can be time-averaged separately because time averaging is a linear operation. The first term is just $2\,T$, twice the kinetic energy.[1] The second term may be rewritten using the force on particle a, $\vec{F}_a$, using Newton's second law. We may thus write the above as

$$\langle T \rangle = -\frac{1}{2} \left\langle \sum_a \vec{F}_a \cdot \vec{r}_a \right\rangle$$

The quantity on the right side is known as the **virial**. The key result is that the time-averaged kinetic energy is related to a time-average of a quantity involving the forces and positions. The virial theorem is similar to the work-energy theorem, which relates the work done by a force on particle to the particle's kinetic energy and which is also derived using Newtons' second law, but the virial theorem pertains to time-averaged, summed quantities rather than to individual particle instantaneous quantities. What good does this do for us? The key is the time-averaging and summing over particles, which lets the virial theorem be used in unexpected ways.

Example: Ideal Gas Law

We can, for example, use the virial theorem to prove the ideal gas law! Consider a gas of temperature Θ confined to a box of volume V. The temperature is defined in terms of the average (over particles) kinetic energy of the gas particles, so we can relate the total time-averaged kinetic energy of the gas to the temperature:

$$\langle T \rangle = \frac{3}{2} N\, k\, \Theta$$

where N is the number of gas particles. To calculate the virial, we need to evlauate the time average of $\vec{F}_a \cdot \vec{r}_a$. The gas particles move freely except when they hit a wall, when an instantaneous force is exerted to reflect them from the wall. Let us write the sum for virial, $\sum_a \vec{F}_a \cdot \vec{r}_a$, as an integral over the walls of the box. The average contribution to the force exerted on an area element dA of the wall by the gas at any instant in time is

$$d\vec{F} = \hat{n}\, P\, dA$$

where $\hat{n}$ is the outward normal at the wall. By Newton's third law, the force exerted on the gas by the wall is the same modulo a sign. The sum for the virial is then just an integral over the walls:

$$-\frac{1}{2} \left\langle \sum_a \vec{F}_a \cdot \vec{r}_a \right\rangle = \frac{1}{2} \int_S P\, \hat{n} \cdot \vec{r}\, dA$$

[1] Note: there is no ambiguity here about how to calculate T. $\vec{p}_a$ and $\vec{r}_a$ are not generalized coordinates, they are the Cartesian vectors describing the particles (think back to elementary mechanics). It always holds that $\vec{p}_a = m_a\, \dot{\vec{r}}_a$ and that $T = \frac{1}{2}\, m_a\, \dot{\vec{r}}_a^2$, hence $\vec{p}_a \cdot \dot{\vec{r}}_a = 2\,T$.

where S indicates the closed surface defining the box walls. We may do the surface integral using Gauss' theorem:

$$\int_S \hat{n} \cdot \vec{r}\, dA = \int_V \vec{\nabla} \cdot \vec{r}\, dV = 3\, V$$

Thus, we obtain

$$\begin{aligned} \langle T \rangle &= \frac{3}{2} P\, V \\ \Longrightarrow \quad N\, k\, \Theta &= P\, V \end{aligned}$$

which is the ideal gas law. Note especially how we did the derivation using only information about the time-averaged force: we didn't need to know any details about the interaction of the particles with the walls except the average force per unit area, P, due to that interaction.

Conservative Power Law Potentials

If we now consider the specific case of particles being acted upon by a conservative force field that is derived from a potential energy that is a power law in particle radius from the center of force, we can evaluate the virial more explicitly. That is, we assume

$$\vec{F}_a = -\vec{\nabla}_a V(\vec{r}_a)$$

where $V(\vec{r})$ is the potential energy and where $\vec{\nabla}_a$ is the gradient with respect to particle a's position vector, $\vec{r}_a$. Note that we are assuming that all the particles move in a single potential energy that is a function of the particle position.[2] This assumption allows us to write the virial as

$$-\frac{1}{2} \left\langle \sum_a \vec{F}_a \cdot \vec{r}_a \right\rangle = \frac{1}{2} \left\langle \sum_a \vec{r}_a \cdot \vec{\nabla}_a V(\vec{r}_a) \right\rangle$$

Now, assume $V(\vec{r}_a) = k\, r_a^n$. Then, $\vec{\nabla}_a V(\vec{r}_a) = n\, k\, r_a^{n-1}\, \hat{r}_a$ and the virial becomes

$$\begin{aligned} -\frac{1}{2} \left\langle \sum_a \vec{F}_a \cdot \vec{r}_a \right\rangle &= \frac{1}{2} \left\langle \sum_a r_a\, n\, k\, r_a^{n-1} \right\rangle = \frac{n}{2} \left\langle \sum_a k\, r_a^n \right\rangle \\ &= \frac{n}{2} \left\langle \sum_a V(r_a) \right\rangle = \frac{n}{2} \langle U \rangle \end{aligned}$$

where $U = \sum_a V(r_a)$ is the total potential energy of the system. Thus, the virial theorem reduces to

$$\langle T \rangle = \frac{n}{2} \langle U \rangle$$

[2] Strictly speaking, pairwise central forces do not satisfy this form. But, for an ensemble of many particles, it is a very good approximation to say that each particle moves in a potential generated by the whole ensemble that looks like a potential fixed to the center of mass of the ensemble, which we take to be at rest. The ensemble potential is quite close to independent of the position of any single particle.

That is, we obtain a very simple relation between the time-averaged kinetic and potential energies of the system.

Example: The Virial Theorem in Astrophysics

The virial theorem is used widely in astrophysics because of the dominance of gravity and because it relates directly observable quantities – kinetic energy and temperature – to unobservable quantities – potential energy and mass. We assume gravitational forces, so $n = -1$. If we divide the virial theorem by the number of particles, we have

$$\frac{1}{N}\langle T\rangle = \frac{1}{2}\frac{1}{N}\left|\langle U\rangle\right|$$

(The sign in n has been canceled by the use of the absolute value sign.) That is, the kinetic energy per particle is half the potential energy per particle. We can use this in different ways to measure total masses of systems:

- If we are looking at a gas cloud, we can measure the gas temperature Θ by its free-free photon emission.[3] That gives us $\langle T\rangle$. The potential energy can be rewritten in terms of the cloud mass M, the typical gas particle mass μ, and the cloud-averaged particle radius. We denote this latter averaged radius as, somewhat uninformatively, the virial radius, R_v. The virial theorem then tells us

$$\frac{3}{2}\,k\,\Theta = \frac{1}{N}\langle T\rangle = \frac{1}{N}\frac{1}{2}\left|\langle U\rangle\right| = \frac{1}{2}\,G\,M\,\mu\left\langle\frac{1}{r}\right\rangle \equiv \frac{1}{2}\,G\,M\,\mu\frac{1}{R_v}$$

$$3\,k\,\Theta = \frac{G\,M\,\mu}{R_v}$$

  Note that the averaging is done on $1/r$, not on r. A typical application would be to use the virial theorem to measure the cloud mass. One has to assume that the cloud is spherically symmetric and optically transparent to its own free-free emission; one can then infer from the observed photon radial distribution the shape (but not the normalization!) of the cloud's density profile. From the shape of unnormalized profile, one can calculate the virial radius. The gas is almost always mostly ionized hydrogen, so μ is known. That leaves the cloud mass as the only unknown. Thus, one can infer from only the photon emission and the virial theorem the cloud mass without any absolute knowledge of normalization of the photon emission in terms of the density. That's rather remarkable!

- If we are looking at a galaxy, we can measure the line-of-sight velocity of a subset of stars by redshift of known spectral lines. The same technique works for galaxies orbiting in a galaxy clusters. Assuming isotropy of the object, the line-of-sight velocity and the velocity transverse to the line of sight will be equal on average (up to a $\sqrt{2}$). Assuming all the orbiting objects in the larger object are of roughly equal mass, the kinetic energy

[3] Free-free emission is just the process of electrons scattering via the Coulomb force off ions in a plasma, a gas that is hot enough that the bulk of the atoms are ionized. Since the electrons are accelerated in these scattering events, they emit light in the form of a photon. The typical photon energy depends on the plasma temperature; for the very hot plasma in galaxy clusters, which is at millions of degrees K, the photons are keV-energy X-rays. In our own galaxy, the emission is usually in the radio, with wavelength of 1 cm and longer.

per particle is simply related to the rms of the measured line-of-sight velocity:

$$\frac{1}{N}\langle T\rangle = \frac{1}{2}\, m\, v_{3d,rms}^2 = \frac{3}{2}\, m\, v_{1d,rms}^2$$

where we relate the full 3-dimensional rms velocity to the measured one-dimension rms velocity assuming isotropy. We can do the same kind of thing as we did for the gas cloud, except now m will drop out:

$$\frac{3}{2}\, m\, v_{1d,rms}^2 = \frac{1}{N}\langle T\rangle = \frac{1}{N}\,\frac{1}{2}\,|\langle U\rangle| = \frac{1}{2}\, G\, M\, m\, \langle 1r\rangle \equiv \frac{1}{2}\, G\, M\, m \frac{1}{R_v}$$

$$3\, v_{1d,rms}^2 = \frac{G\, M}{R_v}$$

Since the test particles whose velocities we measure are discrete objects, we can just make a plot of their number density as a function of radius and from that calculate the virial radius. We can thus determine M only from our observations of the test particle positions and line-of-sight velocities!

1.3.3 Collisions of Particles

Thus section wa formerly section 1.3.2.

Elastic Collisions: Energy

- In the calculation of the kinetic energy in the lab frame, we conside rthe special case $m_1 = m_2$. We incorrectly state

$$\frac{T_1}{T_0} = \frac{4\, m_1^2}{(m_1 + m_2)^2} = 1$$

The expression should be

$$\frac{T_1}{T_0} = \frac{4\, m_1^2}{(m_1 + m_2)^2}\, \cos^2\psi_1 = \cos^2\psi_1$$

Chapter 2

Lagrangian and Hamiltonian Dynamics

2.1 The Lagrangian Approach to Mechanics

2.1.2 Virtual Displacement, Virtual Work, and Generalized Forces

Virtual Work

- Toward the end of the subsection, just before Example 2.1, we have a discussion about how constraint forces satisfy the assumption that they do no virtual work. The discussion could stand to be rewritten somewhat. We replace the original text

> **At this point, we specialize to constraints that do no net work when a virtual displacement is applied.** This assumption is critical. Making this assumption implies that only the *non-constraint* forces need be included in the sum over j because the terms due to constraints yield no contribution. One must go case-by-case in deciding whether a constraint satisfies this assumption.[1]
>
> Mathematically, the assumption lets us drop the part of the sum containing constraint forces, leaving
>
> $$\delta W = \sum_{ij} \vec{F}_{ij}^{(nc)} \cdot \delta \vec{r}_i$$
>
> where the $^{(nc)}$ superscript indicates that the sum is only over non-constraint forces. Note that each constraint force term may or may not vanish separately. If $\vec{F}_{ij} \cdot \delta \vec{r}_i$ vanishes for a single particle i and a single constraint j, then, the constraint force must act on only one particle and must act normal to the motion. Our elliptical wire constraint is of this form. Not all constraints are

[1] It is not clear whether it is possible to state general rules about which kinds of constraints satisfy the assumption. In fact, Schaum's Outline on Lagrangian Dynamics (D. A. Wells) says "While the truth of this statement is easily demonstrated with simple examples, a general proof is usually not attempted. It may be regarded as a postulate." Goldstein simply states that "We now restrict ourselves to systems for which *the net virtual work of the forces of constraint is zero*." and makes no statement about the general applicability of the assumption. Note that Hand and Finch completely gloss over this subtlety; they simply state "Recall that since constraint forces always act to maintain the constraint, they point in a direction perpendicular to the movement of the parts of the system. This means that the constraint forces do not contribute anything to the virtual work." The first sentence is patently false, as our Atwood's machine example shows!

of this form; see our Atwood's machine example below. In my experience, it is true that $\sum_i \vec{F}_{ij} \cdot \delta\vec{r}_i$ always vanishes for any given j: that is, one constraint may require multiple forces acting on multiple particles, but the net work done to enforce that constraint is always zero. The second example below is of this type. But our assumption is more generic; it simply says that

$$\sum_{ij} \vec{F}_{ij}^{(c)} \cdot \delta\vec{r}_i = 0$$

where the $^{(c)}$ superscript restricts the sum to constraint forces but the sum is over *all* constraint forces and *all* particles.

with the new text

At this point, we specialize to constraints that do no net work when a virtual displacement is applied. This assumption is critical. Making this assumption implies that only the *non-constraint* forces need be included in the sum over j because the terms due to constraints yield no contribution.

The assumption deserves some detailed discussion. It is not clear whether it is possible to state general rules about which kinds of constraints satisfy the assumption. In fact, Schaum's Outline on Lagrangian Dynamics (D. A. Wells) says "While the truth of this statement is easily demonstrated with simple examples, a general proof is usually not attempted. It may be regarded as a postulate." Goldstein simply states that "We now restrict ourselves to systems for which *the net virtual work of the forces of constraint is zero*" and makes no statement about the general applicability of the assumption. Note that Hand and Finch completely gloss over this subtlety; they simply state "Recall that since constraint forces always act to maintain the constraint, they point in a direction perpendicular to the movement of the parts of the system. This means that the constraint forces do not contribute anything to the virtual work." The first sentence is patently false, as our Atwood's machine example shows!

Let us try to at least get an intuitive idea of how different kinds of constraints satisfy the assumption. There are clearly three kinds:

1. "normal forces": If $\vec{F}_{ij} \cdot \delta\vec{r}_i$ vanishes for a single particle i and a single constraint j, then, the constraint force must act on only one particle and must act normal to the motion. Our elliptical wire constraint is of this form. The constraint defining rigid-body motion, $|\vec{r}_a - \vec{r}_b| = c_{ab}$ for all particles a, b in the body, is similar in form: an allowed virtual displacement keeps the length of the vector separation of the two particles fixed but allows its orientation to change, while the force that maintains the constraint is the central force between the two, which acts along the separation vector and thus perpendicular to the virtual displacement. It is not quite the same as the single-particle version, but it still can be considered a normal force because the constraint force and virtual displacement are perpendicular.
2. "single-constraint satisfaction": Not all constraints are "normal forces"; see our Atwood's machine example below, where the constraint force acts along the virtual displacement so that $\vec{F}_{ij} \cdot \delta\vec{r}_i \neq 0$ but $\sum_i \vec{F}_{ij} \cdot \delta\vec{r}_i$ does vanish due to summation over i. In this case, once one sums over the particles

that are affected by a particular constraint, then the sum vanishes. For this type of constraint, each constraint j satisfies the assumption $\sum_i \vec{F}_{ij} \cdot \delta\vec{r}_i = 0$ independently. Of course, normal forces are a special subset of this class, but it is instructive to consider them separately.

3. "interlocking constraint satisfaction": I admittedly cannot think of an example, but one can imagine in a general sense that some set of interlocking constraints, where multiple coordinates appear in multiple constraints, might require the summation over both i and j for the assumption to hold.

Because of the possibility that there exist situations of the third type, we use the most generic assumption we need to proceed with our derivation, which is the third one. We write that down as

$$\sum_{ij} \vec{F}_{ij}^{(c)} \cdot \delta\vec{r}_i = 0$$

where the $^{(c)}$ superscript restricts the sum to constraint forces but the sum is over *all* constraint forces and *all* particles. Mathematically, the assumption lets us drop the part of the virtual work sum containing constraint forces, leaving

$$\delta W = \sum_{ij} \vec{F}_{ij}^{(nc)} \cdot \delta\vec{r}_i$$

where the $^{(nc)}$ superscript indicates that the sum is only over non-constraint forces.

2.1.9 Special Nonconservative Cases

Velocity-Dependent Potentials

- In this section, we make use of the vector identity

$$\vec{a} \times (\vec{\nabla} \times \vec{b}) = \vec{\nabla} (\vec{a} \cdot \vec{b}) - (\vec{a} \cdot \vec{\nabla}) \vec{b}$$

We should have stated that this identity only holds when $\vec{a}$ is not acted on by $\vec{\nabla}$: we have moved $\vec{a}$ from the left side of $\vec{\nabla}$ (LHS of equation) to its right side (first term on RHS). More generally, one has to be a bit more careful. It is not possible to clearly write the general result using just vector notation, but it can be written using index notation:

$$\left[\vec{a} \times (\vec{\nabla} \times \vec{b})\right]_i = \sum_j a_j \nabla_i b_j - \sum_j a_j \nabla_j b_i$$

The key point is that in the first term, $\vec{a}$ is in a dot product with $\vec{b}$, but $\vec{\nabla}$ must be allowed to act on $\vec{b}$ first, and not as $\vec{\nabla} \cdot \vec{b}$.

2.2 Variational Calculus and Dynamics

2.2.3 Imposing Constraints in Variational Dynamics

Lagrange Multipliers for Standard Calculus Minimization Problems

- The paragraph that describes why the equation

$$\vec{\nabla} H(\vec{y}^m) = \sum_p \lambda_p(\vec{y}^m) \vec{\nabla} G_p(\vec{y}^m)$$

is the correct requirement on $\vec{\nabla}H$ for minimization of H subject to the constraints given by the $\{G_p\}$ is not as clear as it could be. The original paragraph is

> That is, if there is a point $\vec{y}^m$ such that the vector $\vec{\nabla}H$ at that point is a linear combination of the vectors $\vec{\nabla}G_p$ at that point, then we are assured that $\vec{\nabla}H(\vec{y}^m)\cdot d\vec{y} = 0$ for all $d\vec{y}$ that satisfy the constraints (*i.e.*, for all $d\vec{y}$ that satisfy $\vec{\nabla}G_p(\vec{y}^m)\cdot d\vec{y} = 0$). Geometrically, the constraints require that the allowed $d\vec{y}$ lie in the intersection of the j tangent subspaces created by the constraints. In order to have $dH = 0$ subject to the constraints, the gradient of H must point "out of" this intersection of subspaces – it can have no component "along" the intersection of tangent subspaces. If $\vec{\nabla}H$ is perpendicular to at least one of the tangent subspaces (denote it by $T_\perp$), then $\vec{\nabla}H$ is perpendicular to the intersection of all the tangent subspaces because that intersection contains some portion of $T_\perp$ (if it did not, the intersection would be empty, indicating the constraints are not self-consistent). $\vec{\nabla}H$ is perpendicular to at least one tangent subspace if it is a linear combination of the above form. The undetermined $\{\lambda_p\}$ can be found because we now have M minimization equations (the M components of Equation 2.27) and j constraint equations (the j equations $\{G_p(\vec{y}) - C_p = 0\}$), enough information to find the M components of $\vec{y}^m$ and the j parameters $\{\lambda_p\}$.

Replace it with the following:

> We can see that this expression is sufficient to minimize H while respecting the constraint condition by simply calculating $\vec{\nabla}H(\vec{y}^m)\cdot d\vec{y}$:
>
> $$\vec{\nabla}H(\vec{y}^m)\cdot d\vec{y} = \left[\sum_p \lambda_p(\vec{y}^m)\vec{\nabla}G_p(\vec{y}^m)\right]\cdot d\vec{y} = \sum_p \lambda_p(\vec{y}^m)\left[\vec{\nabla}G_p(\vec{y}^m)\cdot d\vec{y}\right] = 0$$
>
> where the last equality holds because we are only consideringq $d\vec{y}$ that satisfy the constraints, which mathematically is the statement $\vec{\nabla}G_p(\vec{y}^m)\cdot d\vec{y} = 0$ for all p as explained earlier.
>
> How do we see that the expression is necessary, that it is the minimal possible expression for $\vec{\nabla}H(\vec{y}^m)$? That is straightforward to see geometrically. Geometrically, the constraints require that the allowed $d\vec{y}$ lie in the intersection of the j tangent subspaces created by the constraints. In order to have $dH = 0$ subject to the constraints, the gradient of H must point "out of" this intersection of subspaces – it can have no component "along" the intersection of tangent subspaces. To stay out of this intersection, $\vec{\nabla}H(\vec{y}^m)$ must have a nonzero projection along at least one $\vec{\nabla}G_p(\vec{y}^m)$. We can use proof by contradiction to see this. Suppose $\vec{\nabla}H(\vec{y}^m)$ has zero projection along every $\vec{\nabla}G_p(\vec{y}^m)$. Then $\vec{\nabla}H(\vec{y}^m)$ would be perpendicular to all $\vec{\nabla}G_p(\vec{y}^m)$, which would imply that $\vec{\nabla}H(\vec{y}^m)$ lies in all the tangent subspaces, which implies that it lies in the intersection of the tangent subspaces. That is exactly what we do *not* want. So it must be false to assume that $\vec{\nabla}H(\vec{y}^m)$ has zero projection along every $\vec{\nabla}G_p(\vec{y}^m)$. If that is false, then the expression we have written down is the minimal one that allows $\vec{\nabla}H(\vec{y}^m)$ to have nonzero projection along at least one $\vec{\nabla}G_p(\vec{y}^m)$. Not all the λ_p need be nonzero, only one has to be nonzero.[2]

[2] We shall see later that a given λ_p vanishes when the corresponding constraint force vanishes. That happens if no

The undetermined $\{\lambda_p\}$ can be found because we now have M minimization equations (the M components of Equation 2.27) and j constraint equations (the j equations $\{G_p(\vec{y}) - C_p = 0\}$), enough information to find the M components of $\vec{y}^m$ and the j parameters $\{\lambda_p\}$.

Lagrange Multipliers in Variational Problems

- We make the statement

  Applying the variation $\{\delta y_k\}$ to the constraint yields the j equations

  $$\sum_k \frac{\partial G_p}{\partial y_k}\,\delta y_k = 0$$

  It would be better to explicitly show the x dependence in the above equation by writing it as

  Applying the variation $\{\delta y_k(x)\}$ to the constraint yields the j equations

  $$\sum_k \frac{\partial G_p(x)}{\partial y_k}\,\delta y_k(x) = 0 \qquad \text{at each } x \text{ independently}$$

Lagrange Multipliers and Constraint Forces

- **Example 2.9**
  Toward the end of the example, where we say

  Solving, we find

  $$\ddot{x} = \frac{1}{2}\,g\,\sin\alpha \quad \ddot{y} = \frac{1}{2}\,\frac{g\,\sin\alpha}{R} \quad \lambda = -\frac{1}{2}\,M\,g\sin\alpha$$

  the $\ddot{y}$ in the second equation should obviously be a $\ddot{\theta}$. The corrected version is

  Solving, we find

  $$\ddot{x} = \frac{1}{2}\,g\,\sin\alpha \quad \ddot{\theta} = \frac{1}{2}\,\frac{g\,\sin\alpha}{R} \quad \lambda = -\frac{1}{2}\,M\,g\sin\alpha$$

2.2.4 Incorporating Nonholonomic Constraints in Variational Dynamics

We break this section into two subsections, the first one describing an example of how to deal with inequality constraints, the second one being the original discussion of nonintegrable constraints. The first section is completely new, so we provide it here. We also list an error in the old section on nonintegrable constraints.

---

force is needed to enforce the constraint. For example, if a particle is restricted to live on the plane $x = 0$, is subject to gravity in the z direction, and is given initial condition $x = 0$, then it will continue to satisfy $x = 0$ for all time with no constraint force applied.

Inequality Constraints

We give an example of how to deal with one type of inequality constraint. We make no claim this is a generic method, but it is instructive.
Consider a pointlike particle sitting on top of a hemisphere of radius R. Let the coordinate system origin be at the center of the hemisphere. The particle thus satisfies $r \geq R$. Suppose the particle is placed at rest at the top of the hemisphere and given an infinitesimal nudge to get it sliding down the hemisphere. We want to determine the dynamics of the particle, specifically the polar angle at which it leaves the hemisphere. The polar angle is taken to be zero at the top of the hemisphere.
We will solve the problem by treating the constraint as an exact equality constraint, $r = R$, and then finding at what angle the constraint force enforcing the constraint – as given by the corresponding Lagrange multiplier – goes to zero. It will become clear that this is the point at which the particle leaves the hemisphere.
The Lagrangian and constraint equation are

$$L = \frac{1}{2}\, m\, \dot{r}^2 + \frac{1}{2}\, m\, r^2\, \dot{\theta}^2 - m\, g\, r\, \cos\theta \qquad\qquad r - R = 0$$

The resulting Euler-Lagrange equations with Lagrange multipliers are:

$$\begin{aligned} r: &\quad m\, r\, \dot{\theta}^2 - m\, g\, \cos\theta - m\, \ddot{r} + \lambda = 0 \\ \theta: &\quad m\, g\, r\, \sin\theta - \frac{d}{dt}\left(m\, r^2\, \dot{\theta}\right) = 0 \\ \lambda: &\quad r - R = 0 \end{aligned}$$

We use the λ equation to substitute for r (also using $\dot{r} = 0$) to obtain

$$\begin{aligned} r: &\quad \lambda = m\, g\, \cos\theta - m\, R\, \dot{\theta}^2 \\ \theta: &\quad m\, R^2 \ddot{\theta} = m\, g\, r\, \sin\theta \end{aligned}$$

We cannot solve the system analytically. But it is clear that λ tells us the force that the hemisphere must exert to counter gravity acting on the particle, after subtracting off the apparent centrifugal force due to the particle's circular motion. λ is positive when the particle is on the hemisphere, but it will go to zero and become negative when the particle leaves the hemisphere. (The equations obviously become invalid as soon as λ goes negative because the constraint is no longer valid.) So we simply need to find the θ at which $\lambda = 0$. So we want $\dot{\theta}$ as a function of θ. We can obtain that by conservation of energy:

$$\begin{aligned} m\, g\, R &= m\, g\, R\, \cos\theta + \frac{1}{2}\, m\, R^2 \dot{\theta}^2 \\ \dot{\theta}^2 &= 2\, \frac{g}{R}\, (1 - \cos\theta) \end{aligned}$$

We could also obtain the same result by a clever integration of the θ equation of motion:

$$\begin{aligned} \ddot{\theta} &= \frac{d\dot{\theta}}{dt} = \frac{d\dot{\theta}}{d\theta}\, \frac{d\theta}{dt} = \dot{\theta}\, \frac{d\dot{\theta}}{d\theta} \\ \ddot{\theta}\, d\theta &= \dot{\theta}\, d\dot{\theta} \end{aligned}$$

We can replace $\ddot{\theta}$ on the LHS using the θ equation of motion and then integrate both sides:

$$\begin{aligned} \frac{g}{R} \int_0^\theta \sin\theta' d\theta' &= \frac{1}{2}\, \left(\dot{\theta}'\right)^2 \Big|_0^{\dot{\theta}} \\ \frac{g}{R}\, (1 - \cos\theta) &= \frac{1}{2}\, \dot{\theta}^2 \end{aligned}$$

With $\dot{\theta}$ written in terms of θ, we can reduce the r equation and obtain λ as a function of θ:

$$\begin{aligned}\lambda &= m\,g\,\cos\theta - m\,R\left[2\,\frac{g}{R}\,(1-\cos\theta)\right] \\ &= m\,g\,\left(3\,\cos\theta - 2\right)\end{aligned}$$

So $\lambda = 0$ when $\cos\theta = 2/3$. One can obtain the complete dynamics by calculating $\dot{\theta}$ at this angle and then using this position and velocity as the initial condition for continued motion with no constraint and subject to gravity. That would, for example, let one figure out exactly where the particle hits the ground.

Nonintegrable Constraints

- There is an error in the Lagrangian written down for the problem given in the notes. We wrote down

$$L = \frac{1}{2}\,MR^2\left[\frac{3}{2}\,\dot{\phi}^2 + \frac{1}{4}\,\dot{\theta}^2\right]$$

  It should be

$$L = \frac{1}{2}\,MR^2\left[\frac{3}{2}\,\dot{\phi}^2 + \frac{1}{4}\,\dot{\theta}^2\right] - m\,g\,y\,\sin\alpha$$

  The expression we originally gave was just the kinetic energy.

2.3 Hamiltonian Dynamics

2.3.2 Phase Space and Liouville's Theorem

Liouville's Theorem

- At the start, we incorrectly state that Liouville's theorem implies that the phase space fluid is incompressible. Rather, it is a fluid with incompressible flow. This will be explained in more detail in the following revised section.

Alternate Versions of Liouville's Theorem

This section is rather badly muddled and badly named. The title should be changed the **The Fluid Interpretation of Liouville's Theorem**. Just replace it wholesale with the following:

> To see clearly the fluid interpretation of Liouville's theorem, we must expand out the differential above or the total derivative using the chain rule: If we write out the total derivative using the chain rule, we have
>
> $$\rho(\vec{q}(t+dt),\vec{p}(t+dt),t+dt) - \rho(\vec{q}(t),\vec{p}(t),t) = \frac{\partial\rho}{\partial t}\,dt + \sum_k\left[\frac{\partial\rho}{\partial q_k}\,\dot{q_k}\,dt + \frac{\partial\rho}{\partial p_k}\,\dot{p_k}\,dt\right]$$
>
> or
>
> $$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_k\left[\frac{\partial\rho}{\partial q_k}\,\dot{q_k} + \frac{\partial\rho}{\partial p_k}\,\dot{p_k}\right]$$

The second term can be rewritten to give

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \left[\dot{\vec{q}} \cdot \vec{\nabla}_q + \dot{\vec{p}} \cdot \vec{\nabla}_p\right] \rho = \frac{\partial \rho}{\partial t} + \dot{\vec{\xi}} \cdot \vec{\nabla}_\xi \rho$$

where in the first step we have simply rewritten the sums over k as dot products (note the q and p subscripts on $\vec{\nabla}$, which specify which variables the gradient is taken with respect to), and in the second step we have defined a new composite coordinate $\vec{\xi}$, with

$$\xi_k = \begin{cases} q_{(k+1)/2} & k = \text{odd} \\ p_{k/2} & k = \text{even} \end{cases}$$

This is known as *symplectic notation*; we will return to this term later. Four points can be made about this form:

- The derivative expression given is completely analogous to the one we would find in fluid mechanics, where $\dot{\vec{\xi}}$ would be replaced by the fluid velocity $\vec{v}$ and the gradient $\vec{\nabla}_\xi$ would be replaced by the simple spatial gradient $\vec{\nabla}_r$. The derivative operator $\frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla}_r$ is generically called the "convective", "advective", "Lagrangian", "substantial", "substantive", "material", or "Stokes" derivative due to its use in fluid mechanics to calculate differentials "moving along" with the fluid.
- Liouville's theorem can be rewritten in a slightly different way using the convective form. Let's separate the two pieces of the convective derivative:

  $$0 = \frac{d\rho}{dt} \quad \Longrightarrow \quad 0 = \frac{\partial \rho}{\partial t} + \dot{\vec{\xi}} \cdot \vec{\nabla}_\xi \rho \quad \Longrightarrow \quad \frac{\partial \rho}{\partial t} = -\dot{\vec{\xi}} \cdot \vec{\nabla}_\xi \rho$$

  The last form says that if, instead of moving along with the phase space flow, you sit at one point in phase space and watch the phase space density at that point change with time, then the rate at which it changes, $\frac{\partial \rho}{\partial t}$, is given by the negative of the gradient of the phase space density along the flow direction multiplied by the flow speed.
- Liouville's theorem is not a trivial result of conservation of particle number. Conservation of particle number simply states

  $$\frac{\partial}{\partial t} \int_V \rho \, dV_{qp} + \int_S dA_{qp} \; \hat{n} \cdot \rho \, \dot{\vec{\xi}} = 0$$

  i.e., the rate of change of the number of particles in a volume V in phase space is just related to the net flow into the volume through its surface S. Gauss's theorem lets us rewrite the above as

  $$\frac{\partial}{\partial t} \int_V \rho \, dV_{qp} + \int_V dV_{qp} \; \vec{\nabla}_\xi \cdot \left(\rho \, \dot{\vec{\xi}}\right) = 0$$

  Since the volume is not changing in time (we consider a volume V fixed in phase space, not one moving with the flow), we may move the time derivative inside the integral:

  $$\int_V dV_{qp} \left[\frac{\partial \rho}{\partial t} + \vec{\nabla}_\xi \cdot \left(\rho \, \dot{\vec{\xi}}\right)\right] = 0$$

Finally, because the volume V is arbitrary, the integrand must vanish at any point in phase space:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_\xi \cdot \left(\rho \dot{\vec{\xi}} \right) = 0$$

The above equation is the **continuity equation** and simply states conservation of particle number. We need to know nothing about Hamiltonian dynamics to derive it To get from the continuity equation to Liouville's theorem, let's expand out the divergence term:

$$0 = \frac{\partial \rho}{\partial t} + \vec{\nabla}_\xi \cdot \left(\rho \dot{\vec{\xi}} \right) = \frac{\partial \rho}{\partial t} + \dot{\vec{\xi}} \cdot \vec{\nabla}_\xi \rho + \rho \left(\vec{\nabla}_\xi \cdot \dot{\vec{\xi}} \right)$$

We need for the third term to vanish to obtain Liouville's theorem. It turns out that it vanishes because of Hamilton's equations:

$$\vec{\nabla}_\xi \cdot \dot{\vec{\xi}} = \sum_k \left[\frac{\partial \dot{q}_k}{\partial q_k} + \frac{\partial \dot{p}_k}{\partial p_k} \right] = \sum_k \left[\frac{\partial^2 H}{\partial q_k \, \partial p_k} - \frac{\partial^2 H}{\partial p_k \, \partial q_k} \right] = 0$$

Thus, we are left with

$$0 = \frac{\partial \rho}{\partial t} + \dot{\vec{\xi}} \cdot \vec{\nabla}_\xi \rho$$

i.e., Liouville's theorem. Liouville's theorem is thus a consequence of both conservation of particle number and Hamilton's equations. We implicitly assumed conservation of particle number in our derivation: we followed particular particles along, assuming that they could not vanish. And we had to use Hamilton's equations in our original derivation of Liouville's theorem. So we have obtained an alternate derivation of Liouville's theorem, though one that rests on the same physics. The derivation is actually really the same in that one can make a term-by-term correspondence between the two; the language we use here is just more sophisticated and results in a more compact derivation.

- Combining the above two points, we can see that there is a very nice analogy between phase space flow and incompressibility in fluid mechanics. There are two kinds of incompressibility in fluid mechanics, and it is the latter one that corresponds to phase space flow:
  - *incompressible fluid*: An incompressible fluid has an unchangeable density: $\rho =$ constant. Therefore, $\vec{\nabla}_r \rho = 0$, $\frac{\partial \rho}{\partial t} = 0$, and $\frac{d\rho}{dt} = 0$. We can use the continuity equation to also infer that $\vec{\nabla}_r \cdot \vec{v} = 0$:

    $$0 = \frac{\partial \rho}{\partial t} + \vec{\nabla}_r \cdot (\rho \, \vec{v}) = \frac{\partial \rho}{\partial t} + \vec{v} \cdot \vec{\nabla}_r \rho + \rho \, \vec{\nabla}_r \cdot \vec{v}$$

    Incompressibility of the fluid implies that the first two terms vanish, so (assuming we don't have a trivial fluid with $\rho = 0$)

    $$\vec{\nabla}_r \cdot \vec{v} = 0$$

    We thus see than an incompressible fluid is very uninteresting. The phase space density is *not* an incompressible fluid.

– *incompressible flow*: The definition of a fluid with incompressible flow is

$$\vec{\nabla}_r \cdot \vec{v} = 0$$

We derived above that the phase space flow obeys the analogous equation thanks to Hamilton's equations. Thus, there is a perfect correspondence between the phase space fluid and a fluid with incompressible flow. Just as we derived Liouville's theorem from the continuity equation with the additional condition based on Hamilton's equations that $\vec{\nabla}_\xi \cdot \dot{\vec{\xi}} = 0$, we could derive the analogy of Liouville's theorem for incompressible flows:

$$0 = \frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \vec{v} \cdot \vec{\nabla}_r \rho$$

We note than an incompressible fluid always has incompressible flow (because we showed above $\vec{\nabla}_r \cdot \vec{v} = 0$ for an incompressible fluid), but of course a fluid with incompressible flow need not be an incompressible fluid.

2.4 Topics in Theoretical Mechanics

2.4.1 Canonical Transformations and Generating Functions

Generating Functions

- At one point, we state

> We shall make the enlightened guess that a *sufficient* condition for that term to vanish is that the *Lagrangians* (not the Hamiltonians) in the two coordinate systems differ by a total derivative of a function $F(q, Q, t)$ and that $P = \frac{\partial F}{\partial Q}$.

There is a sign error typo on the condition on P, the statement should be

> We shall make the enlightened guess that a *sufficient* condition for that term to vanish is that the *Lagrangians* (not the Hamiltonians) in the two coordinate systems differ by a total derivative of a function $F(q, Q, t)$ and that $P = -\frac{\partial F}{\partial Q}$.

2.4.4 Action-Angle Variables and Adiabatic Invariance

Though this section is part of Section 2.4 that is not required for most students, the concepts of action-angle variables and adiabatic invariance are generically useful and should be understood, even if the derivation cannot be appreciated. To that end, we have added a "Pre-Summary and Cookbook" discussion at the start of this section that both helps to outline the results of the derivation that follows and also summarizes the results for those who have not studied the previous material on canonical transformations and generating functions necessary to understand the derivation. This material appears after the initial example of the SHO and before the full derivation begins.

Pre-Summary and Cookbook

In the following section, we shall generalize the above example. The basic result we will obtain is that, for any 1-dimensional periodic system, we may always define conjugate action and angle variables analogous to P and Q above and that they always obey the following:

- The action variable I is constant and has the value

$$I = \oint p(E,q)\,dq \tag{2.1}$$

  where one obtains $p(E,q)$ by inverting $H(p,q)$ to obtain $p(H,q)$ and then using the fact that energy is conserved for a periodic system, so that H is a constant with value E set by the initial conditions, $E = H(p(t=0), q(t=0))$. The integral is thus explicit for any initial condition (though whether it is analytically integrable depends on the specific problem).

- The angle variable evolves linearly in time at rate ω:

$$\psi = \omega\, t + \psi_0 \quad \text{with} \quad \omega \equiv \frac{\partial H(I)}{\partial I} \tag{2.2}$$

  where $H(I)$ is just the Hamiltonian rewritten in terms of I instead of in terms of p and q. ψ_0 is set by initial conditions.

The full proof will demonstrate that it is always possible to define I and ψ for any periodic system and that they always satisfy the above.

We will also demonstrate that I is an *adiabatic invariant*, which means that, if some normally constant parameter in H is changed sufficiently slowly, then I is constant to first order in the change in the parameter. An example of a "normally constant" parameter is the natural frequency ω in a SHO, or the mass m and spring constant k that combine to give ω. By "sufficiently slowly," we mean that, if the changing parameter is α, the rate of change of α must satisfy $\dot{\alpha}\,T/\alpha \ll 1$ where T is the period of the system. That is, the fractional change in α in one period is small compared to 1. By "constant to first order," we mean that $\dot{I} \propto \dot{\alpha}^2$.

Adiabatic invariance of I is how these concepts prove useful because they provide one quantity, I, that is constant even when parameters in the problem and the total energy are changing. The classic example is a pendulum with a slowly lengthening bob length $l(t)$. If the parameter changes sufficiently slowly, we do not need to solve the full problem with the rheonomic constraint that gives us $l(t)$ to find out how the energy, oscillation amplitude, or maximum speed of the pendulum change with time; we simply use the fact that I is constant to relate the rate of change of E to the rate of change of the bob length l, and then we can obtain the rate of change of the amplitude A or the maximum speed v from the rate of change of E.

The geometric interpretation of all of the above in phase space is also interesting. I, by definition, is the area enclosed in phase space by the phase space orbit and ψ is the angular position of the system on that orbit in the pq plane. Adiabatic invariance has a clever geometric interpretation – even as a parameter in the problem changes, causing the orbit shape to change (*e.g.*, the maximum q and p will change), the area of the orbit is preserved as it changes shape.

Additional corrections and clarifications:

Action-Angle Variables

- In the derivation of action-angle variables, we state

Now, we could use the standard Legendre transformation rules to obtain an explicit formula for ψ:

$$\psi = \frac{\partial \widetilde{W}_2(q,I)}{\partial I} = \frac{\partial}{\partial I}\int p(q,I)\,dq = \int dq\,\frac{\partial p(q,I)}{\partial I}$$

Evaluation of this form requires knowing the full dynamics of the system. There is a way around this: because $\widetilde{W}_2$ generates a canonical transformation, we are assured that Hamilton's equations are satisfied in the (ψ, I) variable set. In particular,

$$\dot{\psi} = \frac{\partial H(\psi,I)}{\partial I}$$

Since H is conserved, we know it should be possible to write H as a function of the one constant of the motion available, I, so we are left with

$$\dot{\psi} = \frac{\partial H(I)}{\partial I} \equiv \omega = \text{constant}$$

This of course implies that ψ is cyclic in the transformed Hamiltonian. Since $\dot{\psi}$ is constant, it holds that

$$\psi = \omega\, t + \psi_0$$

We thus obtain an explicit formula for ψ that is valid independent of the details of the problem. The Legendre transformation formula for ψ would yield this, but one has to perform the integral explicitly using the particular $p(q,I)$ function.

The above is not as clear as it could be. We rewrite as follows:

Now, we could use the standard Legendre transformation rules to obtain an explicit formula for ψ:

$$\psi = \frac{\partial \widetilde{W}_2(q,I)}{\partial I} = \frac{\partial}{\partial I}\int p(q,I)\,dq = \int dq\,\frac{\partial p(q,I)}{\partial I}$$

We shall show below that, after the canonical transformation, H has no ψ dependence and thus is a function of I alone, $H = H(I)$. This would allow us to obtain $p(q,I)$ via H: $H = H(p,q)$ also, so we should be able find $p = p(q,H)$ and therefore $p = p(q,H(I)) = p(q,I)$. We could then do the integral and take $\frac{\partial}{\partial I}$, or alternatively take $\frac{\partial}{\partial I}$ and do the integral, to obtain ψ. But it may not always be possible to do the integral analytically. It turns out there is an easier way to obtain $\psi(t)$ using the fact that (ψ, I) are canonical variables by construction. Hamilton's equations in the new variables are

$$\dot{\psi} = \frac{\partial H(\psi,I)}{\partial I} \qquad \dot{I} = -\frac{\partial H(\psi,I)}{\partial \psi}$$

I is constant by construction. Therefore, $\dot{I} = 0$ and thus $\frac{\partial H}{\partial \psi} = 0$. Therefore, the canonically transformed H has no ψ dependence, $H = H(I)$ only. Consequently,

$\dot{\psi} = \frac{\partial H}{\partial I}$ can depend only on I. Moreover, because I is constant in time, we know $\frac{\partial H}{\partial I}$ is constant in time. Putting this all together gives

$$\dot{\psi} = \frac{\partial H(I)}{\partial I} \equiv \omega = \text{constant}$$

Therefore,

$$\psi = \omega\, t + \psi_0$$

We thus obtain an explicit formula for ψ that is valid independent of the details of the problem. The Legendre transformation formula for ψ would yield this, but one would have to perform the integral explicitly using the particular $p(q, I)$ function.

We belabor one point, the functional dependences of ω. We explicitly noted above that $\omega = \omega(I)$: ω is not constant as a function of I. But ω *is* constant in *time* because I is constant in time. The constancy of ω in time provides the simple evolution of the angle variable ψ. Now, in some cases, ω may indeed be independent of I – for example, in the SHO, $\omega = \sqrt{k/m}$ depends only on the parameters of the problem, not on I, which comes from the initial conditions. But, more generally, ω may be a function of I; one example is the non-small-angle pendulum – the period of the motion (which we shall see below is given by ω) is dependent on the initial conditions.

Adiabatic Invariance

- The derivation completes with the following:

  Recall that $\widetilde{W}$ is periodic in ψ because the motion is periodic in ψ. So we have

$$\begin{aligned} -\left\langle \dot{I} \right\rangle &= \frac{\dot{\alpha}}{2\,\pi} \left[\frac{\partial \widetilde{W}(q, \psi, \pi, \alpha(T))}{\partial \alpha} - \frac{\partial \widetilde{W}(q, \psi, \alpha(0))}{\partial \alpha} \right] \\ &= T\, \frac{\dot{\alpha}^2}{2\,\pi}\, \frac{\partial^2 \widetilde{W}}{\partial\, \psi\, \partial\, \alpha} \approx 0 \end{aligned}$$

  There is an error in the last line, it should be

$$-\left\langle \dot{I} \right\rangle = T\, \frac{\dot{\alpha}^2}{2\,\pi}\, \frac{\partial^2 \widetilde{W}}{\partial \alpha^2} \approx 0$$

2.4.5 The Hamilton-Jacobi Equation

Examples

- **Example 2.16: Simple Harmonic Oscillator**
  At the end of the example, where we calculate S explicitly using the solution for $q(t)$ and $p(t)$ we have obtained, we have the line

$$S = -E\, t + \sqrt{2\, E} \int d\Big(\frac{\sqrt{2\, E}}{\omega}\, \sin\left(\omega\, t + \phi\right)\, \sqrt{1 - \sin^2\left(\omega\, t + \phi\right)}$$

The parentheses clearly do not match up, it should be

$$S = -E\,t + \sqrt{2\,E} \int \left[d\left(\frac{\sqrt{2\,E}}{\omega} \sin\left(\omega\,t + \phi\right) \right) \right] \sqrt{1 - \sin^2\left(\omega\,t + \phi\right)}$$

The d is taking the differential of the quantity $\left(\frac{\sqrt{2\,E}}{\omega} \sin\left(\omega\,t + \phi\right) \right)$.

Chapter 3

Oscillations

3.1 The Simple Harmonic Oscillator

3.1.3 The Damped Simple Harmonic Oscillator

Damped Solutions

- In the discussion of the critically damped case, we derive the form for the first of the two possible solutions. We state

$$\begin{aligned} q_1(t) &= \exp(-t)\left[\exp\left(t\sqrt{\epsilon}\right) + \exp\left(-t\sqrt{\epsilon}\right)\right] \\ &\approx \exp(-t) \end{aligned}$$

  We lost a factor of 2 in stepping between the two lines as we let $\epsilon \to 0$, the last line should be

$$q_1(t) \approx 2\,\exp(-t)$$

  The loss of the factor of 2 has no other repurcussions.

3.1.4 The Driven Simple and Damped Harmonic Oscillator

Delta-Function Response, or Green's Functions

- In the paragraph

> Intuitively, one should think of the Green's function $G(t, t')$ as giving the motion at a time t if there was in impulsive force of unit amplitude at time t'. Since the differential equation is linear, we obviously can find the response to two separate impulses a and b at times t_a and t_b by calculating $q(t) = a\,G(t, t_a) + b\,G(t, t_b)$. For a force consisting of impulses F_i at times t_i, *i.e.*, $F(t) = \sum_i \delta(t - t_i)$, the solution will be $q(t) = \sum_i F_i\,G(t, t_i)$. The extension to a continous force is trivial, giving Equations 3.8 and 3.9 above.

  we give the wrong expression for $F(t)$, it should be

$$F(t) = \sum_i F_i \delta(t - t_i)$$

- In the explicit derivation of the generic initial conditions for the Green's function, we have

The explicit method for finding the Green's function is to make use of the homogeneous solution and the properties of the δ-function. Away from $t = t'$, the Green's function satisfies the homogeneous differential equation. Causality requires $G(t, t') = 0$ for $t < t'$, which is a perfectly good solution of the differential equation. For $t > t'$, we can just take G to be the generic solution to the homogeneous equation. The coefficients A and B are obtained by integrating the differential equation for G (Equation 3.9) over a small interval around $t = t'$ and making careful use of the delta function's properties; essentially, this integration provides boundary conditions for G that determine A and B:

$$\int_{t'-\epsilon}^{t'+\epsilon} dt \left[\ddot{G}(t,t') + \frac{\dot{G}(t,t')}{Q} + G(t,t')\right] = \int_{t'-\epsilon}^{t'+\epsilon} dt\, \delta(t-t') = 1$$

$$\left[\dot{G}(t'+\epsilon,t') - \dot{G}(t'-\epsilon,t')\right] + \frac{1}{Q}\left[G(t'+\epsilon,t') - G(t'-\epsilon,t')\right] + 2\,\epsilon\, G(t',t') = 1$$

$$\dot{G}(t'+\epsilon,t') + \frac{1}{Q}\, G(t'+\epsilon,t') + 2\,\epsilon\, G(t',t') = 1$$

$$\dot{G}(t'+\epsilon,t') + \frac{1}{Q}\left[G(t',t') + \epsilon\,\dot{G}(t'^{+},t')\right] + 2\,\epsilon\, G(t',t') = 1$$

where we have made use of multiple facts in stepping through the above: 1) $G(t, t')$ and $\dot{G}(t, t')$ vanish for $t < t'$; 2) G must be continuous at $t = t'$ to avoid having the oscillator change position discontinously 3) $\dot{G}(t, t')$ may change discontinuously at $t = t'$ because of the impulsive force, but it is finite everywhere, so any terms of the form $\epsilon\,\dot{G}$ vanish as $\epsilon \to 0$. So we have the initial conditions

$$\lim_{\epsilon\to 0}\ \left[G(t',t') + 2\,\epsilon\, G(t',t')\right] = G(t',t') = 0$$

$$\lim_{\epsilon\to 0}\ \left[\dot{G}(t'+\epsilon,t') + \epsilon\,\dot{G}(t'^{+},t')\right] = \dot{G}(t'^{+},t') = 1$$

The above was both unclear and not quite right. We should have done it as follows:

The explicit method for finding the Green's function is to make use of the homogeneous solution and the properties of the δ-function. Away from $t = t'$, the Green's function satisfies the homogeneous differential equation. Causality requires $G(t, t') = 0$ for $t < t'$, which is a perfectly good solution of the differential equation. Continuity of position then implies $G(t', t') = 0$ too (*i.e.*, the position cannot change instantaneously, even with and infinite delta-function force). For $t > t'$, we can just take G to be the generic solution to the homogeneous equation. The coefficients A and B are obtained by integrating the differential equation for G (Equation 3.9) over a small interval around $t = t'$ and making careful use of the delta function's properties; essentially, this integration provides initial conditions for G that determine A and B:

$$\int_{t'-\epsilon}^{t'+\epsilon} dt \left[\ddot{G}(t,t') + \frac{\dot{G}(t,t')}{Q} + G(t,t')\right] = \int_{t'-\epsilon}^{t'+\epsilon} dt\, \delta(t-t')$$

$$\left[\dot{G}(t'+\epsilon,t') - \dot{G}(t'-\epsilon,t')\right] + \frac{1}{Q}\left[G(t'+\epsilon,t') - G(t'-\epsilon,t')\right] + 2\,\epsilon\, G(t',t') = 1$$

where we have explicitly done the integrations. Now, use the facts $\dot{G}(t'-\epsilon, t') = 0$ and $G(t'-\epsilon, t') = 0$ (which result from causality), giving

$$\dot{G}(t'+\epsilon, t') + \frac{1}{Q}\, G(t'+\epsilon, t') + 2\,\epsilon\, G(t', t') = 1$$

Now, Taylor-expand the second term to obtain

$$\dot{G}(t'+\epsilon, t') + \frac{1}{Q}\left[G(t', t') + \epsilon\, \dot{G}(t'^{+}, t')\right] + 2\,\epsilon\, G(t', t') = 1$$

where $\dot{G}(t'^{+}, t')$ is the derivative of G as $t \to 0$ from the positive side. This positive-side/negative-side distinction is necessary for $\dot{G}$ because it may change discontinuously at $t = t'$ (because the acceleration will be infinite due to the delta-function applied force.) Next, use the fact explained above that $G(t', t') = 0$ due to causality and continuity of position, yielding:

$$\dot{G}(t'+\epsilon, t') + \frac{\epsilon}{Q}\, \dot{G}(t'^{+}, t') = 1$$

Finally, use the fact that $\dot{G}$, while it may change discontinuously at $t = t'$, must remain finite at all times – even the infinite acceleration provided by the delta-function impulse will integrate to only a step change in velocity. Thus, the second term goes to zero as ϵ goes to zero, giving

$$\dot{G}(t'^{+}, t') = 1$$

Thus, our initial conditions for G are $G(t', t') = 0$ and $\dot{G}(t'^{+}, t') = 1$. The $^{+}$ superscript simply means that the initial condition for the $t > t'$ solution will be that its initial velocity is 1, while for $t < t'$ it holds that $\dot{G} = 0$. The solution will have a velocity discontinuity at $t = t'$. With these and the known form of the solution of the homogeneous equation, we may determine the Green's function completely.

3.2 Coupled Simpler Harmonic Oscillators

3.2.2 General Method of Solution

Finding the Normal Mode Vectors

- We wrote

  > Let's be a bit more explicit and write the equation for the mode vectors in terms of the **cofactors** of the matrix $-\omega^2\,\mathbf{t} + \mathbf{v}$. Recall that the determinant of a matrix can be calculated by taking the sum over the product of any row or column and its cofactors, with sign flips between adjacent elements:

  This turns out to be a strange way to lead in to the discussion. It would be more clear to say

  > We will solve the above equation using the **cofactors** of the matrix $-\omega^2\,\mathbf{t} + \mathbf{v}$. Recall that cofactors arise in the calculation of the determinant of a matrix: the determinant of a matrix can be calculated by taking the sum over the product of any row or column and its cofactors, with sign flips between adjacent elements:

- A bit further below, we wrote

> Now, when we have a matrix equation of the form $\mathbf{a}\,\vec{f} = 0$, we require $|\mathbf{a}| = 0$ to have a nontrivial solution. This is equivalent to requiring that $\vec{r}_i^T \vec{C}_i^r = 0$ or $\vec{c}_i^T \vec{C}_i^c = 0$. That is, any row or column of the matrix $\mathbf{a}$ and its cofactor vector are orthogonal.

While all true, the above suggests that we needed to prove $|\mathbf{a}| = 0$. We already know $|\mathbf{a}| = 0$ because we used that fact to obtain the normal mode frequencies. The above should be rewritten

> Now that you are reminded of what a cofactor is, and we have introduced how the determinant can be written as a dot product of row or column vectors and their respective cofactor vectors, let us note a useful property of these dot products. For the specific $\mathbf{a}$ we are considering here, $\mathbf{a} = \omega_i^2 \mathbf{t} + \mathbf{v}$, we know that $|\mathbf{a}| = 0$: this was necessary to have nontrivial normal mode vectors. In terms of the row/column vectors and their cofactor vectors, this statement implies that $\vec{r}_i^T \vec{C}_i^r = 0$ or $\vec{c}_i^T \vec{C}_i^c = 0$; that is, any row or column of the matrix $\mathbf{a}$ and its cofactor vector are orthogonal. We shall use this orthogonality below.

- There is a typo in the paragraph that follows. The statement

> We can see in fact that any of these cofactor vectors satisfies the full set of equations by realizing that $\sum a_{ij}\,\vec{C}_k^r$ for $k \neq i$ also must vanish because this quantity is the determinant of the matrix obtained by replacing row k of $\mathbf{a}$ with row i and calculating the determinant using row i. The determinant of a matrix with two equal rows vanishes.

should read

> We can see in fact that any of these cofactor vectors satisfies the full set of equations by realizing that $\sum a_{ij} \left(\vec{C}_k^r\right)_j$ for $k \neq i$ also must vanish: this quantity is the determinant of the matrix obtained by replacing row k of $\mathbf{a}$ with row i and calculating the determinant using the new row k (which is now identical to row i); since the determinant of a matrix with two equal rows vanishes, it holds that $\sum a_{ij} \left(\vec{C}_k^r\right)_j$ vanishes even when $k \neq i$.

3.2.4 Degeneracy

General Considerations

- In the first paragraph of this section, wherever we have $M - n$, it should be $M - n - 1$.

Example: Masses Coupled by Springs on a Circle

- We state

> Correctly normalize, the $\lambda = 1$ mode vector is

There is no $\lambda = 1$ mode, this should of course be the $\lambda = 0$ mode.

Chapter 4

Central Force Motion and Scattering

4.2 The Special Case of Gravity – The Kepler Problem

4.2.1 The Shape of Solutions of the Kepler Problem

The General Solution

- When we write out the energy in terms of A, l_θ, and the constants in the problem, the last line is erroneously given as

$$E = \frac{l_\theta^2}{2\,\mu}\left[A^2 + \left(\frac{G\,\mu^2 M}{l_\theta^2}\right)^2\right]$$

  The sign is a typo. The line should be

$$E = \frac{l_\theta^2}{2\,\mu}\left[A^2 - \left(\frac{G\,\mu^2 M}{l_\theta^2}\right)^2\right]$$

  The error only occured here, the remainder of the section is correct.

4.3 Scattering Cross Sections

4.3.2 The Generic Cross Section

Differential Cross Section

- The text introducing the definition of differential cross section is unclear. Replace

  > If dN is the number of particles per unit time scattered into the solid angle $d\Omega$, the differential cross section is
  >
  > $$\frac{d\sigma}{d\Omega}\,d\Omega = \frac{dN}{F}$$
  > $$\frac{d\sigma}{d\Omega} = \frac{1}{F}\,\frac{dN}{d\Omega}$$
  >
  > $\frac{d\sigma}{d\Omega}$ has units of area per steradian; hence the name *cross section*.

  with

If $dN(\theta_*, \phi_*)$ is the number of particles per unit time scattered into the solid angle $d\Omega$ at (θ_*, ϕ_*), we define the differential cross section via the relation

$$\frac{1}{A}\frac{d\sigma}{d\Omega}(\theta_*, \phi_*)\, d\Omega = \frac{dN(\theta_*, \phi_*)}{F\, A}$$

Let us explain the above. On the right side of first line, the denominator is the number of particles per unit time incident on the target from a beam of flux F and cross-sectional area A. Since the numerator is the number of particles per unit time that scatter into $d\Omega$ at (θ_*, ϕ_*), the right side is thus the fraction of particles that scatter into $d\Omega$ at (θ_*, ϕ_*); it is a probability. We include additional factors so that $\frac{d\sigma}{d\Omega}$ is defined only by the scattering force, not by parameters of the experiment. We have a $1/F$ on the right side but none on the left side because including the $1/F$ makes the ratio dN/F independent of F: if F goes up, dN goes up proportionally. However, we include a $1/A$ on the left side to cancel the $1/A$ on the right side because $dN(\theta_*, \phi_*)$ may not scale with A: if one adds cross-sectional area at a radius from which particles do not scatter into the particular solid angle $d\Omega$ at (θ_*, ϕ_*), $dN(\theta_*, \phi_*)$ will not increase when A increases. Hence the different treatment of F and A. Solving for the differential cross section gives

$$\frac{d\sigma}{d\Omega}(\theta_*, \phi_*) = \frac{1}{F}\frac{dN(\theta_*, \phi_*)}{d\Omega}$$

The beam area A has dropped out. $\frac{d\sigma}{d\Omega}$ has units of area per steradian; hence the name *cross section*.

4.3.3 $\frac{1}{r}$ Potentials

Calculating the Differential Cross Section

- There is a simple typo in the result, we wrote

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin\theta_*}\left|\frac{db}{d\theta_*}\right| = \frac{b^2}{4}\csc^2\frac{\theta_*}{2}\sec^2\frac{\theta_*}{2} = \left(\frac{G\,\mu\,M}{4E}\right)^2 \frac{1}{\sin^2\frac{\theta_*}{2}}$$

There is an error in the exponent in the last expression, the last expression should be

$$\left(\frac{G\,\mu\,M}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta_*}{2}}$$

Chapter 5

Rotating Systems

5.2 Dynamics in Rotating Coordinate Systems

5.2.1 Newton's Second Law in Rotating Coordinate Systems

Position and Velocity in Rotating and Non-Rotating Frames

- Following the calculation of $\frac{d}{dt}\,\underline{\vec{r}}\,'(t)$, we have the paragraph

  > The velocity relative to the non-rotating system has two components – one due to rotation of the rotating system, the other due to motion relative to the rotating system. This is not yet what we need, as the right side uses the coordinate representation in the rotating system $\underline{\vec{r}}$. We may use $\underline{\vec{r}}(t) = [\mathbf{R}(t)]^T\,\underline{\vec{r}}\,'(t)$ to rewrite $\underline{\vec{r}}$ in terms of the non-rotating system representation. $\frac{d}{dt}\,\underline{\vec{r}}$ is a bit more tricky. Transforming it by $\mathbf{R}^T$ does not yield $\frac{d}{dt}\,\underline{\vec{r}}\,'$; that would obviously yield a trivial result. What is meant by $\mathbf{R}(t)\,\frac{d}{dt}\,\underline{\vec{r}}(t)$ and why is it not just $\frac{d}{dt}\,\underline{\vec{r}}\,'(t)$?

  This paragraph is confusing because it is not clear at this point why it is undesirable to write the right side in terms of the rotating-system coordinate representation. That point is not important. Replace the paragraph with the following, which focuses on the main point we want to bring up here:

  > The velocity relative to the non-rotating system has two components – one due to rotation of the rotating system, the other due to motion relative to the rotating system. But, what is meant by $\mathbf{R}(t)\,\frac{d}{dt}\,\underline{\vec{r}}(t)$ and why is it not just $\frac{d}{dt}\,\underline{\vec{r}}\,'(t)$?

Rewriting Using Instantaneous Angular Velocity

- In our discussion of the direct method of calculation $\dot{\mathbf{R}}$, we write

$$\begin{aligned}\dot{\mathbf{R}}(t) &= \frac{1}{dt}\left[\mathbf{R}(t+dt) - \mathbf{R}(t)\right] \\ &= \frac{1}{dt}\left[\left(\mathbf{I} + \vec{\omega}(t)\,dt\cdot\vec{\mathbf{M}}\right) - \mathbf{I}\right]\mathbf{R}(t) \\ &= \vec{\omega}(t)\cdot\vec{\mathbf{M}}\,\mathbf{R}(t)\end{aligned}$$

It would be more accurate to write

$$\begin{aligned}\dot{\mathbf{R}}(t) &= \lim_{dt\to 0}\frac{1}{dt}\left[\mathbf{R}(t+dt)-\mathbf{R}(t)\right]\\ &= \lim_{dt\to 0}\frac{1}{dt}\left[\left(\mathbf{I}+\vec{\omega}(t)\,dt\cdot\vec{\mathbf{M}}\right)-\mathbf{I}\right]\mathbf{R}(t)\\ &= \lim_{dt\to 0}\vec{\omega}(t)\cdot\vec{\mathbf{M}}\,\mathbf{R}(t)=\vec{\omega}(t)\cdot\vec{\mathbf{M}}\,\mathbf{R}(t)\end{aligned}$$

5.2.2 Applications

Foucault's Pendulum

- We had a Freudian slip and wrote everywhere "centripetal" instead of "centrifugal" force in this example. The fictitious force is centrifugal, while the true force needed to cause circular motion (or the true force in the rotating frame needed to counter centrifugal force) is called centripetal force.

5.2.3 Lagrangian and Hamiltonian Dynamics in Rotating Coordinate Systems

- In the introduction to this section, we say it is surprising that one can do Lagrangian and Hamiltonian dynamics in rotating coordinate systems. It really is not surprising: the rotating coordinates are just an example of generalized coordinates.

Obtaining and Using the Lagrangians

- We left out some important terms in our calculation of the Euler-Lagrange equations. We wrote

$$\begin{aligned}\frac{d}{dt}\left(m\,\dot{x}-y\,\omega\,\sin\lambda\right) &= -m\,g\,\frac{x}{l} \quad\Longrightarrow\quad \ddot{x}+\frac{g}{l}\,x = 2\,\dot{y}\,\omega\,\sin\lambda\\ \frac{d}{dt}\left(m\,\dot{y}+x\,\omega\,\sin\lambda\right) &= -m\,g\,\frac{y}{l} \quad\Longrightarrow\quad \ddot{y}+\frac{g}{l}\,y = -2\,\dot{x}\,\omega\sin\lambda\end{aligned}$$

  We forgot to include the $\frac{\partial L}{\partial x}$ and $\frac{\partial L}{\partial y}$ terms that come from the extra ω-dependent term in the kinetic energy. The Euler-Lagrange equations should be

$$\begin{aligned}\frac{d}{dt}\left(m\,\dot{x}-y\,\omega\,\sin\lambda\right) &= \dot{y}\,\omega\,\sin\lambda - m\,g\,\frac{x}{l} \qquad\Longrightarrow\qquad \ddot{x}+\frac{g}{l}\,x = 2\,\dot{y}\,\omega\,\sin\lambda\\ \frac{d}{dt}\left(m\,\dot{y}+x\,\omega\,\sin\lambda\right) &= -\dot{x}\,\omega\,\sin\lambda - m\,g\,\frac{y}{l} \qquad\Longrightarrow\qquad \ddot{y}+\frac{g}{l}\,y = -2\,\dot{x}\,\omega\sin\lambda\end{aligned}$$

5.3 Rotational Dynamics of Rigid Bodies

5.3.1 Basic Formalism

Principal Axes

- In the explanation of the coordinate representations of the inertia tensor and the principal axes vectors in an arbitrary non-principal-axis frame F', we write

$$e'_{i,j} = R_{ji} = \left(\mathbf{R}\,\vec{e}_i\right)_j$$

This is wrong because it does not make sense to have $\mathbf{R}$ act on the vector $\vec{e}_i$; it must act on a coordinate representation. The above should be

$$e'_{i,j} = R_{ji} = R_{jk}\delta_{ki} = R_{jk}e_{i,k} = (\mathbf{R}\,\underline{\vec{e}}_i)_j$$

- Similarly, in the explanation of the coordinate representations in the principal-axis frame F, we write

$$e_{i,j} = \delta_{ji} = \left(\mathbf{R}^T\vec{e}_i\right)_j$$

when we should write

$$e_{i,j} = \delta_{ji} = R_{kj}R_{ki} = R_{kj}e'_{i,k} = \left(\mathbf{R}^T\underline{\vec{e}}_i\right)_j$$

Physical Significance of the Moment of Inertia Tensor

- The last paragraph of the section includes the sentence

> In general, we would again need six numbers to fully describe this "mass tensor" – the orientation of the principal axes (axes that yield angular momentum aligned with angular velocity) and the conversion factors from speed to momentum for these three principal axes.

Hopefully, it is clear from context that the use of "angular" in the parenthetical clause is incorrect. The sentence should be

> In general, we would again need six numbers to fully describe this "mass tensor" – the orientation of the principal axes (axes that yield linear momentum aligned with linear velocity) and the conversion factors from speed to momentum for these three principal axes.

Relation of Euler Form to Single-Axis Rotation

- At the start of this section, we say

> We to some extent expect this result – we proved earlier that, for a rigid body, all motion about the center of mass consists of rotation about a common angular velocity vector $\vec{\omega}$ – but this result is more generic, being true for arbitrary rotation (orthogonal) matrices in $N = 3$ dimensions.

In the above, we mistakenly equate rotation matrices with orthogonal matrices. Not all orthogonal matrices are rotation matrices – orthogonal matrices also include coordinate inversions, which cannot in general be written in terms of rotation matrices. So "orthogonal" should be removed from the above statement.

- One step in the proof given in this section involves proving that orthogonal matrices have unit determinant, which is written as follows:

> The defining relation of orthogonal matrices is
>
> $$\mathbf{R}\,\mathbf{R}^T = \mathbf{1}$$

Taking the determinant of both sides, using the facts $|\mathbf{A}\,\mathbf{B}| = |\mathbf{A}||\mathbf{B}|$ and $|\mathbf{A}^T| = |\mathbf{A}|$, we obtain

$$|\mathbf{R}| = |\mathbf{R}^T| = 1$$

This is equivalent to the statement that orthogonal matrices preserve the norm of vectors they operate on.

The last step is not quite correct. It should read

... we obtain

$$|\mathbf{R}| = |\mathbf{R}^T| = \pm 1$$

Now, all rotation matrices must have only the + sign. This can be seen quickly as follows. From our discussion of infinitesimal rotations in terms of the generator matrices $\vec{\mathbf{M}}$, one can see explicitly that an infinitesimal rotation about any single coordinate axis has positive determinant (to linear order in the angle $\delta\theta$ – the corrections are $\mathcal{O}(\delta\theta)^2$), and thus determinant +1. Since infinitesimal rotations commute, one can construct an infinitesimal rotation about any direction from the product of three such rotations; such an infinitesimal rotation must therefore also have determinant +1. It holds that any finite rotation can be built up as an infinite product of infinitesimal rotations, and also that the determinant of a product is the product of the determinants. Since each infinitesimal rotation has determinant +1 to linear order, the product must also have determinant +1 to linear order. Hence, any rotation matrix must have determinant +1.

5.3.2 Torque-Free Motion

Generic Torque-Free Motion: Euler's Equations

- This section needs both an introduction and a conclusion. The introduction should be

  Our goal in this section is to obtain enough information to give the time evolution of the orientation of the rigid body in the space frame. We shall do this by first obtaining the time evolution of the body-frame representation of the angular velocity vector, $\mathbf{R}^T\vec{\underline{\omega}}'(t)$, and then obtaining the rigid body orientation in the space frame from it.

  The conclusion should be

  With the full time evolution of $\mathbf{R}^T\vec{\underline{\omega}}'$ in hand, we can now obtain the orientation of the body as a function of time. First, it holds that

  $$\mathbf{R}^T\vec{\underline{L}}' = \underline{\mathcal{I}}\,\mathbf{R}^T\vec{\underline{\omega}}'$$

  If we apply $\mathbf{R}$ to the above, we obtain

  $$\mathbf{R}\left(\underline{\mathcal{I}}\,\mathbf{R}^T\vec{\underline{\omega}}'\right) = \vec{\underline{L}}'$$

  Now, recall that Euler's equations have given us $\mathbf{R}^T\vec{\underline{\omega}}'$, that $\underline{\mathcal{I}}$ is known because it is the body-frame representation and is thus given by the principal moments,

that $\underline{\vec{L}}'$ is a constant set by the initial conditions, and also that rotation matrices only have three degrees of freedom. Thus, we have a set of three linear equations for the three degrees of freedom of $\mathbf{R}$ at any instant of time t. We can thus obtain $\mathbf{R}(t)$, which completely defines the orientation of the body relative to the space frame as a function of time. Because a rigid body is fully specified by its three principal moments and the orientation of its principal axes, it holds that $\mathbf{R}(t)$, along with the principal moments, fully determines the rigid body in the space frame at any time.

Appendix A

Mathematical Appendix

A.3 Vector and Tensor Definitions and Algebraic Identities

Algebraic Identities

- After the listing of the identities, we make the statement:

> Note that, if any of the vectors are the gradient vector $\vec{\nabla}$, care must be taken in how the above expressions are written out to ensure $\vec{\nabla}$ acts on the appropriate vectors. Any two quantities that commute in the above should be commuted as necessary to get reasonable behavior of $\vec{\nabla}$. One example is the second line, where $\vec{b}$ may need to be moved to the left of $\vec{a} \cdot \vec{c}$ if $\vec{b} = \vec{\nabla}$.

The second sentence is not quite right because in some cases it may not be possible to write the correct form in pure vector notation. Replace the above with

> Note that, if any of the vectors are the gradient vector $\vec{\nabla}$, care must be taken in how the above expressions are written out to ensure $\vec{\nabla}$ acts on the appropriate vectors. Any two quantities that commute in the above should be commuted as necessary to get reasonable behavior of $\vec{\nabla}$. But in some cases, even that may not be sufficient and you will have to keep track of which vector should be acted on by $\vec{\nabla}$. A good example is the second line when $\vec{b} = \vec{\nabla}$. In the simple case of $\vec{a}$ being constant, one simply needs to move the $\vec{b}$ in the first term:
>
> $$\vec{a} \times \left(\vec{\nabla} \times \vec{c}\right) = \vec{\nabla} \left(\vec{a} \cdot \vec{c}\right) - \left(\vec{a} \cdot \vec{\nabla}\right) \vec{c}$$
>
> But if $\vec{a}$ depends on position and does not give zero when acted on by $\vec{\nabla}$, then the above must be read with care. One has to somehow remember that $\vec{\nabla}$ should not be allowed to act on $\vec{a}$ since it does not act on $\vec{a}$ in the original expression. Since the above expression does not correctly convey that meaning, it is better to abandon the vector notation. The completely unambiguous way to write it, using index notation, is
>
> $$\left[\vec{a} \times (\vec{\nabla} \times \vec{c})\right]_i = \sum_j a_j \nabla_i c_j - \sum_j a_j \nabla_j c_i$$
>
> The key point is that in the first term, $\vec{a}$ is in a dot product with $\vec{c}$, but $\vec{\nabla}$ must be allowed to act on $\vec{c}$ first, and not as $\vec{\nabla} \cdot \vec{c}$.

A.6 Taylor Expansion

- The expansion given for $(1 \pm x)^{-1}$ is just wrong. Instead of

$$\frac{1}{1 \pm x} = 1 \mp x - 2\,x^2 + \cdots$$

it should be

$$\frac{1}{1 \pm x} = 1 \mp x - x^2 + \cdots$$

- There are errors in the Taylor expansion of $|\vec{R} - \vec{r}|^n$ also. The corrected version of the derivation is

$$\begin{aligned}
\frac{1}{\left|\vec{R} - \vec{r}\right|^n} &= \frac{1}{\left(R^2 + r^2 - 2\,\vec{R}\cdot\vec{r}\right)^{n/2}} = \frac{1}{R}\,\frac{1}{\left(1 - 2\left(\hat{R}\cdot\hat{r}\right)\frac{r}{R} + \frac{r^2}{R^2}\right)^{n/2}} \\
&= \frac{1}{R}\,\frac{1}{1 + \frac{n}{2}\left[-2\left(\hat{R}\cdot\hat{r}\right)\frac{r}{R} + \frac{r^2}{R^2}\right] + \frac{1}{2}\frac{n}{2}\frac{n-2}{2}\left[-2\left(\hat{R}\cdot\hat{r}\right)\frac{r}{R} + \frac{r^2}{R^2}\right]^2 + \cdots} \\
&= \frac{1}{R}\,\frac{1}{1 + \frac{n}{2}\left[-2\left(\hat{R}\cdot\hat{r}\right)\frac{r}{R} + \frac{r^2}{R^2}\right] + \frac{n(n-2)}{8}\left[4\left(\hat{R}\cdot\hat{r}\right)^2\frac{r^2}{R^2}\right] + \cdots} \\
&= \frac{1}{R}\,\frac{1}{1 - n\left(\hat{R}\cdot\hat{r}\right)\frac{r}{R} + \frac{n}{2}\frac{r^2}{R^2}\left[1 + (n-2)\left(\hat{R}\cdot\hat{r}\right)^2\right] + \cdots} \\
&= \frac{1}{R}\left\{1 + n\,\frac{r}{R}\left(\hat{R}\cdot\hat{r}\right) - \frac{n}{2}\,\frac{r^2}{R^2}\left[1 + (n-2)\left(\hat{R}\cdot\hat{r}\right)^2\right] - n^2\left(\hat{R}\cdot\hat{r}\right)^2\frac{r^2}{R^2} + \cdots\right\} \\
&= \frac{1}{R}\left\{1 + n\,\frac{r}{R}\left(\hat{R}\cdot\hat{r}\right) + \frac{n}{2}\,\frac{r^2}{R^2}\left[-1 - [(n-2) - 2]\left(\hat{R}\cdot\hat{r}\right)^2\right] + \cdots\right\} \\
&= \frac{1}{R}\left\{1 + n\,\frac{r}{R}\left(\hat{R}\cdot\hat{r}\right) + \frac{n}{2}\,\frac{r^2}{R^2}\left[-1 - (n-4)\left(\hat{R}\cdot\hat{r}\right)^2\right] + \cdots\right\}
\end{aligned}$$

For the specific case usually under consideration, $n = 1$, we find

$$\frac{1}{\left|\vec{R} - \vec{r}\right|} = \frac{1}{R}\left\{1 + \frac{r}{R}\left(\hat{R}\cdot\hat{r}\right) + \frac{1}{2}\,\frac{r^2}{R^2}\left[3\left(\hat{R}\cdot\hat{r}\right)^2 - 1\right] + \cdots\right\}$$

Appendix B

Summary of Physical Results Appendix

B.2 Lagrangian and Hamiltonian Dynamics

Theoretical Mechanics

Liouville's Theorem

- Along with the change we made to the section in the main text on this, we should eliminate the statement that an equivalent alternate form is Equation B.65,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_\xi \cdot \left(\dot{\vec{\xi}}\rho\right) = 0$$

  This we explained was simply the statement of conservation of particles. To obtain Liouville's equation from this continuity equation requires use of Hamilton's equations. The continuity equation is therefore not physically the same as Liouville's theorem, the latter has more physical content because it incorporates Hamilton's equations.

Action-Angle Variables and Adiabatic Invariance

- This section should be moved prior to the section on the Hamilton-Jacobi equation. Also, we state here that ψ can be obtained by (Equation B.95)

$$\psi = \frac{\partial}{\partial I} \int p(q, I)\, dq$$

  This is exactly how we said *not* to obtain ψ in our derivation. Replace the entire section with the following:

  > Given a 1-dimensional system whose Hamiltonian is conserved and that undergoes periodic motion, it is possible to define via canonical transformation **action-angle variables** (ψ, I). The action variable is given by
  >
> $$I = \frac{1}{2\pi} \oint p\, dq$$
  >
  > In practice, to get the function $p(q, E)$ in order to do the integral, one obtains from $H(p, q)$ the function $p(q, H)$ and then uses the fact that energy is conserved for periodic systems to replace H by its constant value E. I will depend on E.

The canonical transformation, combined with the fact that H is conserved, allows the Hamiltonian to be rewritten in terms of I. ψ evolves linearly with time according to

$$\psi(t) = \omega\, t + \psi_0 \qquad \text{with} \qquad \omega \equiv \frac{\partial H(I)}{\partial I}$$

ω is constant in time but may depend on I. ω gives the period of the motion,

$$T = \frac{2\,\pi}{\omega}$$

ψ advances by $2\,\pi$ when the motion goes through one period. **Adiabatic invariance** says that if the Hamiltonian has slow time dependence via a parameter α, then the action variable I is to first order independent of changes in H that are slow compared to the period of the periodic motion. This property can be used determine features of the evolution of such systems.

B.4 Central Forces and Dynamics of Scattering

The Kepler Problem

- As we did in the text, we made an error in the energy relation; the first step of Equation B.190 is given as

$$E = \frac{l_\theta^2}{2\,\mu}\left[A^2 + \left(\frac{G\,\mu^2 M}{l_\theta^2}\right)^2\right]$$

It should be

$$E = \frac{l_\theta^2}{2\,\mu}\left[A^2 - \left(\frac{G\,\mu^2 M}{l_\theta^2}\right)^2\right]$$

Dynamics of Scattering

- As we did in the text, we gave the incorrect final result for the Rutherford scattering differential cross section; we gave (Equation B.205)

$$\frac{d\sigma}{d\Omega} = \left(\frac{k}{4E}\right)^2 \frac{1}{\sin^2\frac{\theta_*}{2}}$$

The exponent in the denominator is wrong, the result should of course be

$$\frac{d\sigma}{d\Omega} = \left(\frac{k}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta_*}{2}}$$