Classical Mechanics Meta Review **Herbert Goldstein** **1950** 8.20.08

My raw notes on this book fill 104 pages (the book is 372 pages) and are very detailed. Here I try to summarize each chapter in a general sense, or clarify things that were unclear in my original notes (sometimes at length, for complex chapters). This "meta review" is about 27 pages long. I read this book in the period May 1 → Aug 20, 2008, but over half that time I could not work, so maybe it took 1.7 months to give this excellent book its due after it sat waiting for 38 years.

Contents:

Chapter 1. Survey of the Elementary Principles (1) 1
Chapter 2: Variational Principles and Lagrange's Equations (30). 3
Chapter 3. The two-body central force problem (58) 4
Chapter 4. The Kinematics of Rigid Body Motion (93) 6
Chapter 5. The Rigid Body Equations of Motion (p 143) 7
Chapter 6. Special Relativity in Classical Mechanics (p 185) 14
Chapter 7. The Hamilton Equations of Motion (p 215) 14
Chapter 8. Canonical Transformations (p 237) 15
Chapter 9. Hamilton-Jacobi Theory (p 273) 17
Chapter 10. Small Oscillations (p 318) 25
Chapter 11. Continuous Systems and Fields (p 347) 28

Chapter 1. Survey of the Elementary Principles (1)

The main thrust of this chapter is the derivation of the Lagrange Equations from Newton's Law, importantly including the treatment of "constraints". This is really the tricky aspect of things.

At first, we imagine we have a problem with perhaps 3N coordinates $\mathbf{r}_i$ but we have k constraints of the form $f_\ell(\mathbf{r}_i, t) = 0$ with $\ell = 1..k$. These are not differential constraints involving $d\mathbf{r}_i$, they are "holonomic" constraints involving the whole coordinates $\mathbf{r}_i$. The idea then is that somehow the problem ought to have 3N-k degrees of freedom. Rather than take the independent coordinates as a subset of the $\mathbf{r}_i$, we take them as some "generalized coordinates q_j with j = 1..(3N-k) in terms of which we can express all 3N $\mathbf{r}_i$.

The notion of a virtual displacement $\delta\mathbf{r}_i$ is described as a sort of development tool. With this tool, we can show that $\Sigma_i \mathbf{F}^{(a)}{}_i \bullet \delta\mathbf{r}_i = 0$ in an equilibrium (static) situation, where the forces $\mathbf{F}^{(a)}{}_i$ include only the "applied forces" and we can ignore the forces due to the constraints (which he calls $\mathbf{f}_i$). This result is not quite obvious and is called the principle of virtual work. It *is* obvious if you use the total force, because then we know that $\mathbf{F}^{tot}{}_i = 0$ for each i. It assumes there are no frictional forces.

In a very similar vein, we start with Newton's law and turn the same wheel and we end up with the result that $\Sigma_i (\mathbf{F}^{(a)}{}_i - \dot{\mathbf{p}}_i) \bullet \delta\mathbf{r}_i = 0$. This now applies dynamically, and again it is not quite obvious since we have thrown out the forces of constraint which must be included in Newton's Law $\mathbf{F}_i = \dot{\mathbf{p}}_i$. This line was pursued by D'Alembert and the result is known as his principle.

By making use of the assumed holonomic constraints and the associated generalized coordinates q_j which allow us to say $\mathbf{r}_i = \mathbf{r}_i(..q_j.., t)$, we can derive from **D'Alembert's Principle** the first form of the

Lagrange Equations shown in (1-50) which involves kinetic energy T and shows on the RHS the generalized (applied) forces Q_i which are related pretty trivially to the real applied forces $\mathbf{F}_i$, where he has now dropped the (a) superscript for "applied". As in the Principle, we are allowed to ignore any forces of constraint in using these equations. In the case that the forces $\mathbf{F}_i$ are derived from a potential $V(..\mathbf{r}_i..,t)$, we get the second form (1-53) which involves L = T-V and which has 0 for the RHS. Now no forces appear at all!

In a diversion, we see how we can modify the situation to handle E&M forces which have a violating term linear in velocity, and separately we see how to handle a linear friction situation which is also a violation of our assumption early on that $\Sigma_i \mathbf{f}_i \bullet \delta\mathbf{r}_i = 0$.

Some examples are then presented.

Attempt Better Summary of Chapter 1

1.1 Mechanics of a Particle. Define **p** and **F,** and **L** and **N**, and their Newton's laws. Define work integral as integral of **F** • **dr**, define conservative force and potential V so **F** = -∇V. Show T+V conserved.

1.2 Mechanics of a System of Particles. State the required $\mathbf{F}_{ij}$ rule. Define **R** as **center of mass** position. Show $M\ddot{\mathbf{R}} = \mathbf{F}_{ext}$ and $\mathbf{P} = M\dot{\mathbf{R}}$. Then show $\mathbf{L} = \mathbf{R} \times M\mathbf{v} + \Sigma_i \mathbf{r'}_i \times \mathbf{p'}_i$ (1-26) where prime means relative to CMS point, and interpret as two contributions. Similarly, show $T = (1/2)Mv^2 + \Sigma_i(1/2)m_i v'^2_i$ (1-29). Finally get a parallel statement that $V = \Sigma_i V_i + (1/2)\,\Sigma_{i\neq j} V_{ij}$ where V_{ij} related to F_{ij} and V_i is for external force. For a rigid body, second term is constant -- internal forces do no work. For such a system of particles, you can separate out the CMS motion versus the detailed internal motions.

1.3 Constraints. Newton says $m\mathbf{a} = \mathbf{F}_i^{(e)} + \Sigma_{j\neq i}\mathbf{F}_{ji}$ with those same internal forces. Define a **holonomic** constraint. **Sclero** (no t) and **Rheo** (yes t). Then (1-36) shows idea of **generalized coordinates** so there are then 3N-k of these if you have k constraints. Then **non-holonomic** p 13. Such problems generally need custom solutions. You have 3N-k "**degrees of freedom**" left due to the k constraints.

1.4 D'Alembert's Principle and the Lagrange Equations Start with statics so of course $\Sigma_i \mathbf{F}_i \bullet \delta\mathbf{r}_i = 0$ where each $\mathbf{F}_i = 0$ due to static. Here is where the vague "virtual displacement" first appears p 14. Split $\mathbf{F}_i$ into constraint plus applied, then if constraint part has no work, get $\Sigma_i \mathbf{F}_i^{(a)} \bullet \delta\mathbf{r}_i = 0$ as well. Then page 15 switch from statics to dynamics by replacing $\mathbf{F}_i$ by $\mathbf{F}_i - \dot{\mathbf{p}}_i$ where $\mathbf{F}_i$ is total force. This gives **D'Alembert** as (1-42) where only $\mathbf{F}_i^{(a)}$ appears, big deal is that constraint forces do NOT appear. In my App D I am clearer on all this stuff I think. The no-constraint-work is just an assumption.

Now comes a major set of actions. (1) define the usual q_i generalized coords so $\mathbf{r}_i = \mathbf{r}_i(q_j)$ with same m = 3N-k idea. (2) express $\delta\mathbf{r}_k$ in terms of δq_n . (3) Produce **Lagrange equations** (1-50). Then if conservative, obtain **Euler-Lagrange equations** (1-53), but not called such at this time! I do this in full detail in my Appendix D of Lagrange doc. Lots of support math here that I do better I think.

1.5 Velocity-dependent potentials and the dissipation function. The claim is that IF you can write the generalized force Q_i in the form (1-54) [meaning that a U exists so that is true], THEN you can rescue the Euler Lagrange equations by replacing V in L = T-V with this U object. Above (1-60) G shows that you can in fact do exactly this for the **Lorentz force** and it turns out $U = q\varphi - (q/c)\,\mathbf{A}\bullet\mathbf{v}$. In this case you

do the normal Euler-Lagrange work using L = T-U = T - qφ + (q/c) **A•v** as in (1.61). The problem in this Lorentz force situation is that velocity appears in the Lorentz force. You could not write this force in terms of a velocity-independent potential since then velocity would not appear in the force.

The next application also involves a case where velocity appears in the force, simple **linear friction**. The fix here is that you rescue the Euler Lagrange equations by defining a **certain $\mathcal{F}$ function** (which looks like kinetic energy with $m_i \rightarrow k_i$) and then you just add $\partial\mathcal{F}/\partial\dot{q}_j$ as on page 22 into the E-L equations. This $\mathcal{F}$ is in fact half the rate of heat generation (work done) by friction. Nice bail out.

1.6 Simple Applications of the Lagrangian formulation. When you convert T from Cartesian to some generalized coordinates, if the transformation has no time, so $\partial \mathbf{r}_i(q_j,t)/\partial t = 0$, then only quadratic terms remain in T so T = "**homogeneous quadratic form** in the q_i", not necessarily diagonal. All the following examples have quadratic forms for T.

Here are the **examples**" (1a) motion of one particle Cartesian -- Newton's Law recovered; (1b) motion of one particle polar -- Newton's Law in polars and appearance of ang mom; (2) Two weights on a pulley (**Atwood**) -- trivial, constraint force T never appears. (3) **Bead** on rotating wire. Nothing said about "inertial frames" in this chapter.

As of 10.29.16 I am pretty happy with Goldstein Chapter 1. Wrote much of it into Lag doc.

Chapter 2: Variational Principles and Lagrange's Equations (30).

In the previous chapter we showed that Newton's Law ↔ D'Alembert's Principle ↔ Lagrange Equations. In this chapter, we show that δI(L dt) = 0 ↔ Lagrange Equations. The integral I(Ldt) is *not* called the action here, but I think most people refer to it that way, more later on the book on this. The hypothesis that δI(L dt) = 0 is called **Hamilton's Principle**, and applies if forces are derived from a potential V, ie, if we have a conservative system. When we take a purely mathematical view of the "**variational calculus**" we end up in general with a set of equations that remind us of the Lagrange equations, but they are in the math world called the **Euler-Lagrange equations**. If we then take the special case of t as the variable time, and generalized coordinates q_i as the independent coordinates, then the math Euler-Lagrange equations become the Lagrange Equations of the last chapter, ie, of mechanics.

On page 39 Hamilton's Principle is "extended" to apply to **non-conservative systems** where there is no Lagrangian L, In this case, instead of T-V, you put T+W in the integral, where $W = \Sigma_i \mathbf{F}_i \bullet \mathbf{r}_i$. This leads to the (1-53)-style Lagrange equations.

At this point we consider the following somewhat complicated situation. We have a problem with 3N coordinates to start, we have k holonomic constraints, and we then have n = 3N-k generalized coordinates to deal with. Now on top of this we assume we have another set of m differential (non-holonomic) constraints of the form (2-22) which are linear. This leads us to a set of (1-50)-style Lagrange equations (2-30) which have a non-zero right hand side! This extra stuff involves the "undetermined **Lagrange multipliers**" λ_i and the coefficients in those differential constraints. We realize that this stuff is just the "generalized forces of constraint" Q'_j which we ignored in our Chapter 1 development where we assumed all holonomic constraints. [see Lag App D]

The **generalized momenta** for conserved systems are given by $p_j = \delta L/\delta\dot{q}_j$. This chapter wraps up talking about conservation of total p, of total L, and of energy.

The last section comments that you can apply this analysis to non-mechanics systems such as RLC **circuits**.

Chapter 3. The two-body central force problem (58).

This entire chapter is a simple application of the Lagrange equations with two generalized coordinates r and θ. We have one "particle" which is a reduced mass object μ having coordinate **r**. We can interpret everything as two objects m1 and m2 having the reduced mass μ, and $\mathbf{r} = \mathbf{r_2}\text{-}\mathbf{r_1}$, but we basically solve this as a single particle problem. We assume a central force so potential of form V(r).

Because θ does not appear in V or T, hence not in L, it is **cyclic** and its momentum is conserved. That is of course angular momentum $\boldsymbol{\ell}$ so that is a "quantum number" (one of the 4 integration constants) of the problem, and orbits must be planar because $\mathbf{r}\bullet\mathbf{L} = 0$. The θ Lagrange equation is trivially integrated to get $\boldsymbol{\ell} = \mu r^2\dot{\theta}$. The r Lagrange equation is (3-12) which can be integrated to get (3-15) which has another constant E = T+V, the energy. The r equation can be integrated a second time to get a closed-form result for t(r), the "inverse" r(t), and this is 3-18. Similarly, 3-20 gives t(θ) which one could then solve for θ(t), and then entire problem is solved for **r**(t).

We can include the "centrifugal term" shown in 3-12 in an effective f '(r) [the central force]. We can then examine the "nature" of orbits by looking at the shape of the corresponding V'(r) curve. For a given energy E and angular momentum $\boldsymbol{\ell}$ you can tell if there are closed and/or open orbits. For closed, there will be bounding distances, and for open there will be a closest approach. We then digress a bit into the "**virial theorem**" for a general system, then look at it for a conservative one with a potential, and then a power law central force, and finally for the inverse square law force f(r).

The next major subject is how to find details of the orbit shape, where we don't care so much about time, so we don't care about r(t) or θ(t), but rather we might like so see θ(r). It is convenient to use u = 1/r in place of r, and we get the integral (3-37) for θ(u) for general V(r). Power law potentials are easiest to handle, and some powers give complete solutions in terms of standard special functions. You usually get the so-called "elliptic integrals".

Next, we assume the main case of interest (celestial mechanics, classical Bohr atom) $V = -k/r^2$ and, when the dust settles, the orbits are given by the trivial polar conic section result 3-47 which says $1/r = C(1 + \varepsilon \cos\theta)$ where $\varepsilon = \varepsilon(E,\boldsymbol{\ell}, \mu, k)$ = eccentricity, and C may be related to semi-major axis "a" as in 3-51. The orbit shape classification is then stated in terms of the value of E, always a conic section. All of Kepler's Laws are quickly derived: the #2 equal area in equal time applies for any V(r), while the #3 $\tau^3 \sim a^2$ is only for inverse square, as is the #1 law of conic shapes. Good work for a guy in 1620!

The last section looks more at this same central force problem as a *scattering problem* as done by Rutherford's people where we have an electrostatic inverse square force between projectile +2 alpha particles from a radon source and +N nuclei in thin sheet metal atoms. The CMS cross section is derived to be (3-68) which is $\sin^{-4}(\Theta/2)$ where Θ is the angle of deflection (scattering angle). This is CMS frame because we assume a rigid force center that does not move in the derivation. We finally transform this to a lab frame where the experiment is done. The relation between the scattering angle in the two frames is simply stated by 3-72 for arbitrary m1 and m2, but the cross section requires a derivative as shown in 3-73 and things are a real mess, though doable. For equal masses, there is a simple result. In practice the metal nuclei are very heavy relative to the projectiles so the CMS result is more or less correct. This is an example of purely elastic scattering.

Attempt Better Summary of Chapter 3

3.1 Reduction to 1-body problem. This chapter is about the "effective 1-body problem" where you get r,θ orbits of the 1-body relative to a fixed origin. Gravity problems (like sun-earth, or binary stars) of course have two bodies, and you have to project out the 2-body motions from the 1-body solution. Goldstein alludes to this, but I do the details in "binary star orbits.doc" including Maple plots.

3.2 The Equations of Motion and the First Integrals. The variables for the 1-body effective problem are r,θ. Due to $\boldsymbol{\ell}$ conserved, you know at once that θ is cyclic (does not appear in L), so $\boldsymbol{\ell} = mr^2\dot{\theta}$ is a constant of the motion. This is the first Lagrange equation. This is a "first integral" equation because it concerns $\dot{\theta}$ instead of $\ddot{\theta}$. We later have to do a "second integral" to get to θ from $\dot{\theta}$. Since $dA/dt = (1/2)r^2\dot{\theta}$, this little $\boldsymbol{\ell} = mr^2\dot{\theta}$ gives the "equal area in equal time" **Kepler Law #2**. The second Lagrange Equation for r gives ODE (3-10) and then (3-12) which is an ODE in r which includes $\ddot{r}$ and r and $\boldsymbol{\ell}$, assuming a central potential and $f(r) \equiv -\partial_r V(r)$. By fiddling, Goldstein ends up with an equation only involving $\dot{r}$ which is (3.15), really just energy conservation, and this is then another "first integral". This equation can be integrated a second time to give a complete formal solution (3-18) for t(r) which invert later to get r(t). You can then integrate $\boldsymbol{\ell} = mr(t)^2\dot{\theta}$ to get a formal integral for θ(t) as in (3-20). These two are the "second integrals" and we then have a complete formal solution to the 1-body effective problem. Aside: We know $\dot{r} = v_r$ completely just from (3-16), and we know v completely from (3-21), so then we also know v_θ. This is all "first integral" information and we know it all without solving the problem!

3.3 Classification of orbits for the effective 1-body problem. This is a good graphical section showing things like turning points for various potential shapes. It is like QM in that you have "bound state" orbits and "unbound" ones. I have written up the complaint about Fig 3-7 which cannot be correct for any purely attractive potential and which appears on the 3rd edition cover! But authors bail it out as possible for a linear combination of attractive and repulsive potentials. The energy E of your 1-body system determines the nature of the orbit. The point is that you learn a lot about orbits even before you solve the problem for a given potential.

3.4 The Virial Theorem. This slightly out-of-place discussion shows how to relate <V> to <T> where these are long-term time averages for the orbiting effective 1-body. For a simple power potential you find $\langle T\rangle = (1/2)(n+1)\langle V\rangle$ and then for the inverse-square law n = -2 you get the famous result $\langle T\rangle = -(1/2)\langle V\rangle$. Here $V = \alpha r^{n+1}$. For inverse square this is V = -k/r with n = -2 and V is always negative, so <V> is negative. I have not tried to translate things like this to the 2-body real problem.

3.5 The orbit ODE and which potential powers are integrable. Two important facts here : (1) dt and dθ are trivially related as in (3-31); (2) people tend to use u = 1/r as the radial variable. This gives a little ODE in variable θ for function u(θ) as in (3-34). Orbits have a certain symmetry. We are now going to do those "second integrals" mentioned at the end of 3.1, but changing dt to dθ. This gives (3-36) as an integral you can try to do to determine θ(r), or (3-37) to get θ(u). You then somehow invert to get r(θ) or u(θ). Then you just let θ run 0 to 2π and you see your *orbit*. G then shows that you can do the integral for

a large range of powers n. For certain powers you get trig/hyper type functions, for other powers you get elliptic integral functions. An interesting topic to ponder, but we really want to get back the planets!

3.6 The Kepler Problem with inverse-square law force. G first rewrites our already-known ODE in u as (3-42) which he converts to p 76 α with its simple solution β where at this point b is unknown. But here is the first view that the orbits are ellipses (for bound orbits) and this is **Kepler Law #1**. G then redoes this derivation by directly integrating our "second integral" for θ which is now (3-44) and he gets the same ellipse equation but now with constants determined and that final result is (3-46). This is the page where I found the benign "Euler substitution" sign error and ended up writing Euler doc. This ellipse formula really does ellipses, parabola and hyperbola where the latter are unbound orbits as shown p 78 α. Circular orbit is a simple case with high-school force balance page 79 α. G shows that semi "a" of orbit depends only on energy E, a = -k/(2E) which is (3-52).

A totally trivial integral of dA/dt over one period τ to get A results in A = $\ell\tau/(2m)$. But A = πab = $\pi a^{3/2}\sqrt{\ell^2/mk}$ as I show in loose-bolt note in raw doc. This then leads to the statement that $\ell\tau/(2m) = \pi a^{3/2}\sqrt{\ell^2/mk}$ and thus $\tau = \pi a^{3/2}\sqrt{4m/k} = 2\pi a^{3/2}\sqrt{m/k}$ which is **Kepler #3** as stated in (3-54). Here a and m are for the effective 1-body problem (see binary stars doc). An approx result is given in (3-56) since sun is so massive. Note that the area law #2 is valid for any central potential, but the other two laws are specific for inverse-square.

3.7 Scattering in a central force field. gaga

Chapter 4. The Kinematics of Rigid Body Motion (93)

We learn that a rigid body has 6 degrees of freedom. We learn about 3x3 real orthogonal rotation matrices he calls A, whose matrix elements we can interpret as direction cosines. Rotations do not commute. With **Euler rotations** $A = R_{\mathbf{z}}(\psi)R_{\mathbf{x}}(\theta)R_{\mathbf{z}}(\phi)$, the 3x3 passive matrix is as shown in 4-46. In the 2D representation which I call SU(2), the four complex matrix elements are called **Cayley-Klein parameters** and are of course just functions of ψ, θ and ϕ as shown in 4-67 (passive). You can relate these worlds by saying that $P' = QPQ^{\dagger}$ where $P = \mathbf{r}\bullet\boldsymbol{\sigma}$ and Q in SU(2).

Euler's Theorem says that any finite rotation can be represented as an axis $\hat{\mathbf{n}}$ and an angle γ, and **Chasle's Theorem** says the most general displacement of a rigid body is a rotation plus a translation (Galilean group element). [G does not mention that these commute.]

Next, with some help from me, we have development of the following rules:

$$[d\mathbf{G}]_{\mathbf{space}} = [d\mathbf{G}]_{\mathbf{body}} + d\boldsymbol{\Omega} \times \mathbf{G} \qquad \text{where } d\boldsymbol{\Omega} = \gamma\,\hat{\mathbf{n}}$$

$$[d\mathbf{G}/dt]_{\mathbf{space}} = [d\mathbf{G}/dt]_{\mathbf{body}} + \boldsymbol{\omega} \times \mathbf{G} \qquad \text{where } \boldsymbol{\omega} = d\boldsymbol{\Omega}/dt \qquad 4\text{-}100$$

with γ being a differentially small angle. **G** is an arbitrary vector quantity. I use subscripts to show the observing frame, and superscripts to show the frame in which coordinates are expressed. If the superscript is omitted, you can use either frame to express the coordinates.

$X^{\mathbf{body}}$ = the quantity X "as expressed in body frame coordinates".
$X_{\mathbf{body}}$ = the quantity X "as measured in the body frame"

At this point, G digresses to write the vector **ω** the body frame coordinates, and we find that

$(\dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi, \dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi, \dot{\phi}\cos\theta + \dot{\psi}) = \boldsymbol{\omega}^{\mathbf{body}} = (x', y', z')$ components

and the derivation is shown in my full notes in full detail. This is an example where we should use the superscript "body".

We then apply our general rule twice to find that

$$m\mathbf{a}_{\mathbf{space}} = m\mathbf{a}_{\mathbf{body}} + 2m\boldsymbol{\omega} \times \mathbf{v}_{\mathbf{body}} + m\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r}$$

$$\mathbf{F}_{\mathbf{space}} = \mathbf{F}_{\mathbf{body}} + 2m\boldsymbol{\omega} \times \mathbf{v}_{\mathbf{body}} + m\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r}$$

$$\mathbf{F}_{\mathbf{body}} = \mathbf{F}_{\mathbf{space}} - 2m\boldsymbol{\omega} \times \mathbf{v}_{\mathbf{body}} - m\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r} = \mathbf{F}_{\mathbf{space}} + \text{Coriolis} + \text{Centrifugal}$$

where **a** = acceleration, **v** = velocity, **F** = force. If you work in the body frame, then, you see the true space forces, plus you see two extra "fictitious" forces both caused by the fact that your body frame is rotating, and these are the **Coriolis** and **Centrifugal** forces. In the NH, the Coriolis force pushes particles to the right (unless trajectory is exactly aligned with latitude), reverse in SH. This results in cyclones, ballistic corrections for both horizontal and vertical motions, and precessing pendulums, as well as fancy Hertzberg interactions between vibrational and rotational atomic energy levels.

The above ω x G type rule will play major roles in the next chapter.

7Chapter 5. The Rigid Body Equations of Motion (p 143)

This chapter's meta review is very long because a lot happens here and things are complicated. Remember that Klein and Sommerfeld write four huge volumes on the Theory of Tops!

In section 5.1, G derives the rigid body fact that

$\mathbf{L} = I\boldsymbol{\omega}$ // analogous to $\mathbf{p} = m\mathbf{v}$

where I is the **inertia *tensor*** (whereas mass m is a scalar),

$$I_{jk} = \int dV\, \rho(\mathbf{r})\,(r^2\delta_{jk} - r_j r_k)$$

The key fact in this derivation is that $\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i$ as observed in the space frame [$\mathbf{L} = \mathbf{r} \times m\mathbf{v}$] , and thus we make use of the little theorem of the previous chapter. Right off the bat, then, we see that **L** and **ω** are generally *not* in the same direction. Only if all three moments were equal would they be in the same direction, as for a sphere or cube, *or* if **ω** happens to lie along a principle axis. For example, if $\boldsymbol{\omega} = \omega_z\hat{\mathbf{z}}$ and z is a principle axis, then $\mathbf{L} = I_3\boldsymbol{\omega}$.

In section 5.2, we get off on the subject of **dyads and dyadics**. When the dust settles, you realize that the **dyad** [AB] is just a particular class of square matrix constructed from two vectors, and we can write [AB] = AB^T where we think of A on the right as a non-square matrix that is a column vector, and B^T similarly but as a row vector. Each side of this equation is a 2D matrix. In particular we have $[AB]_{ij} = A_iB_j$. You can express each of the terms in the inertia tensor as a dyad $[I] = \Sigma_i m_i \ (r_i^2 [1] - [r_i r_i])$ where [1] is of course just the identity matrix. A linear combination of dyads is called a **dyadic**, so [I] is a dyadic. A typical notation you see is something like **a•I•b** where **a** and **b** are vectors and **I** is a dyadic. My notation would be **a**•[I]•**b** to show with square brackets and I is meant as a dyadic. The idea seems to be that I (unbolded) is a matrix, but **I** bolded is a dyadic. Unfortunately, G also uses I to represent a certain I = **n•I•n**. But since he doesn't use the matrix I, I guess that is OK. In my mind, we have

a•I•b = a•[I]•b = **a** • I**b** = a^T I b = a scalar number, // = <a| I |b> in bra-ket notation

where a^T for example is a row vector. Similarly we might see something like this

L = **I•ω** = [I]•ω = I **ω**

where the dyadic notation appears to be a dot product between a matrix and a vector.

Meanwhile, [1] = [$\hat{x}\hat{x}$] + [$\hat{y}\hat{y}$] + [$\hat{z}\hat{z}$] is in fact a dyadic, not a dyad, since it is the sum of dyads. Think of [$\hat{x}\hat{x}$] as a 3x3 matrix with a 1 in the upper left corner. Usually this appears as just 1 = $\hat{x}\hat{x}$ + $\hat{y}\hat{y}$ + $\hat{z}\hat{z}$ and of course you can say $[AB]_{ij} = A_1B_1[\hat{x}\hat{x}]$ + 8 other terms.

All in all, this notation adds nothing at all to clarity in my opinion.

For matrices with this special form, it is true that you can use the following little rules:

[AB]•C = AB^TC = A(B^TC) = A (B•C)
C•[AB] = C^TA B^T = (C•A) B^T
A•[BC]•D = A^T BC^T D = A^TB C^TD = (A•B)(C•D)

where the first two rows are vectors (second is a row vector) and the last a scalar. As you can see, using normal matrix notation gives the right answers in each case. I'll bet later editions skip dyadics!

One small detail is this: once you pick a set of body axes for your rigid body, the inertia tensor I is a constant in time, just the way mass is a constant. Usually you pick principle body axes to get the I tensor to be diagonal, as we see in the next section.

In section 5.3, we obtain this new result (again, thinking ½ $m_i v_i^2$ = ½ m_i (**ω** x $\mathbf{r}_i$)^2)

T = ½ **ω•I•ω** = ½ ω^T I ω in our various notations

Since **ω** = ω**n** we get

T = ½ ω^2 **n•I•n** ≡ ½ I_n ω^2 . // analogous to T = ½ mv^2

So the item analogous to linear mass m is the inertia moment $I_n \equiv \mathbf{n \bullet I \bullet n}$,

$I_n \equiv \mathbf{n \bullet I \bullet n} = \Sigma_i m_i \{ r_i^2 - (r_i \bullet n)^2 \} \equiv$ the **moment of inertia** *about* rotation axis **n**

but G writes this as just I, not I_n, which is a little confusing since matrix I is also I.

Therefore, since $\mathbf{L} = I\,\boldsymbol{\omega}$ we can write

$T = \frac{1}{2}\,\boldsymbol{\omega} \bullet \mathbf{L}$ // analogous to $T = \frac{1}{2}\,\mathbf{v} \bullet \mathbf{p} = \frac{1}{2}\,\mathbf{v} \bullet m\mathbf{v} = \frac{1}{2}\,mv^2$

G shows us that the above definition of I_n is the same as the usual definition of using the perpendicular distance squared from the rotation axis to weight each mass element,

$I_n = \mathbf{n \bullet I \bullet n} = \Sigma_i m_i\,(\mathbf{r}_i \times \mathbf{n})^2$

Then we have the "**parallel axis theorem**" which says $I_a = I_c + Md^2$. Here, I_c is an any axis passing through the CMS, and I_a is another axis parallel to I_c distance d away, and of course M is the total mass of the rigid body.

In section 5.4 G shows in more detail than I need in my current life that since I is symmetric and real, it can be **diagonalized** by transforming to the principle axes of the rigid body. For me, this is just a matrix theorem on my list of many theorems. Yes, need GSO choices if two diagonal moments equal.

The second item of interest is to consider again $T = \frac{1}{2}\,\boldsymbol{\omega} \bullet \mathbf{I} \bullet \boldsymbol{\omega} = \frac{1}{2}\,\omega^{\mathbf{T}} I\,\omega$, so that $2T = I_{11}\omega_x^2 + I_{12}\,\omega_x\omega_y + \ldots$ where I assume non-principle coordinate axes for the moment. This equation describes a weirdly oriented **inertia ellipsoid** in ω-space, and the vector **ω** must lie somewhere on the surface of this ellipsoid. If you write this in terms of vector $\mathbf{n} = \boldsymbol{\omega}/\omega$, in place of 2T you get I_n for the ellipse in n-space. And if you write in terms of $\boldsymbol{\rho} = \mathbf{n}/\sqrt{I_n}$,you get 1 for the ellipse in ρ-space. Compare $x^2 + y^2 = R^2$ to $x^2 + y^2 = 1$ for a normalized circle.

Finally, we define the **radius of gyration** $\mathbf{R}_0$ for a given axis n as $I_n = MR_o^2$ which obviously has the dimensions of distance.

In section 5.5, we write the principle-axis Lagrangian L = T-V and then start to write down the 1-50 form Lagrange equations which involve only T, not L (so we can ignore V). Our generalized coordinates are taken to be the **Euler angles** which describe the rigid body's position, and we assume a potential V as a function of these coordinates, so L = T-V top page 157. G writes down the Lagrange equation for angle ψ since this is "aligned" with one of the body's principle axes which we used to call z' but we now call z. In writing L in 5-31, we assumed we were using principle axes remember! The generalized force corresponding to ψ is N_z (in the body coordinates, but not measured in the body rotating frame!). This Lagrange equation comes out like so:

$I_3\dot{\omega}_z - \omega_x\omega_y(I_1 - I_2) = N_z$ + cyclic = 5-34

where I could put things like N_z^{body}, but we don't bother. Instead of then looking at the other two Lagrange equations, we recognize the above single equation as one of three components of the following vector equation,

$I\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (I\boldsymbol{\omega}) = \mathbf{N}$ "**Euler's equations** of motion for a rigid body"

and then we can express the various vectors in any frame we want (but in the body frame the matrix I will be diagonal). We can obtain the other two component equations just doing "cyclic" to get 5-34. The other two component equations are NOT the other two Lagrange equations because θ and φ are not associated with principle axes (rather with space z and line of nodes), so this set of equations (or the one vector equation) gets the special name "Euler's equations of motion for a rigid body".

The variable of interest here is $\boldsymbol{\omega}(t)$. If we can find $\boldsymbol{\omega}(t)$ (remember, it moves on the ellipse in ω-space) then our problem is solved. We then know that $\mathbf{L} = I\boldsymbol{\omega}$ and $T = \frac{1}{2}\,\boldsymbol{\omega}\bullet I\bullet\boldsymbol{\omega}$, for example.

In a linear mechanics problem we want to find $\mathbf{r}(t)$ for a particle, from which we can get $\mathbf{v}(t)$. Yes, we want x,y,z in the linear, and here we want φ,θ,ψ, but it happens that the equation for $\boldsymbol{\omega}(t)$ is relatively simple. Maybe the right comparison is to compare the above to $m\mathbf{a} = \mathbf{F}$ which you want to solve for $\mathbf{a}(t)$ and then integrate twice.

Remember from Chapter 4 we had:

$$\boldsymbol{\omega}^{body} = (\dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi,\ \dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi,\ \dot{\phi}\cos\theta + \dot{\psi})$$

which is then three equations

$$\omega_x = \dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi$$
$$\omega_y = \dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi$$
$$\omega_z = \dot{\phi}\cos\theta + \dot{\psi}$$

If we can solve the Euler equation for $\boldsymbol{\omega}(t)$ in the body frame, then we can perhaps use the above three equations to determine the angles like θ(t). These equations are pretty messy, though, since they involve angle rates linearly, but angles trigonometrically!

In section 5.6 we set external torque $\mathbf{N} = 0$ and we think about "**force and torque free motion" of a rigid body**. We don't care about the CMS translation, so just assume CMS is fixed in space. So our Euler equations are now

$I_3\dot{\omega}_z - \omega_x\omega_y(I_1 - I_2) = 0$ + cyclic = 5-34

$I\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (I\boldsymbol{\omega}) = 0$

Our first foray in this subject is to describe the **Poinsot Construction**. Recall that $\boldsymbol{\omega}(t)$ must lie on the ellipsoid which does not vary in scale because T is constant (conserved, no torques). Because **L** is also

conserved, G shows that the solution **ω**(t) goes to a point on the ellipsoid which touches (at a single point) a so-called invariable plane whose distance from the ellipsoid center is a constant 2T/L. With the ellipsoid fixed, once could think of **ω**(t) as moving on the ellipsoid and the invariant plane then swings around keeping its distance from the center constant. This is a pretty severe constraint on the motion of **ω**(t) and gives a qualitative description of the solution. Usually one thinks of the plane as fixed, and the ellipsoid then **rolls without slipping** on the plane. As this happens, **ω**(t) varies, being the vector from the fixed ellipsoid origin to the point of contact. The path of **ω**(t) on the plane is called the **herpolhode**, and on the ellipsoid is called the **polhode**, names coined by Poinsot. The motion of the rolling ellipsoid then gives you the motion of your rigid body. Though the two are not the same thing, they are aligned. The **ω**(t) vector describes some weird circulating path on the invariant plane, and you then have some sort of **generalized precession** of the tip of the **ω** vector. If the object has $I_1 = I_2$, this herpolhode path is a circle, and you have what we normally call precession.

Case where $I_1 = I_2$: At this point, bottom of page 161, G sets out to solve the problem in the $I_1 = I_2$ axial case. We quickly find that ω_z = constant and then the Euler equations describe the usual precessing vector and we find out that the **precession frequency** is $\Omega = (1-I_3/I_1)\,\omega_z$ which gets slower as the object becomes "rounder" in terms of its moments. My favorite example is a **football** spinning on its axis at ω_z and $\boldsymbol{\omega}_T$ then does this precession. Notice that for this symmetrical object, there is no nutation! G points out that the **earth** itself would precess in this manner were it a true rigid object with a 300 day period since it is slightly oblate, but in reality the planet is not rigid and measurement gives more like 427 days. You can throw a football such that it only does the spiral and the precession amplitude is zero, so whether or not there actually is precession depends on the "initial conditions". The Poinsot picture is an axial ellipse rolling on the plane and the herpolhode is a circle. The ellipse rolls around the axis between center and plane. I presume one could solve the general problem with three different moments, but suspect elliptic functions then appear.

In section 5-7 we attack the **"heavy symmetrical top"**, an example where there is an external force and torque. The top has mass M, gravity is g, distance from cms to contact point is $\boldsymbol{\ell}$, moments are $I_2 = I_1$ and I_3. In the first part of the analysis below, the "heavy" assumption is not used.

Presumably this could be attacked using the Euler equations above, but instead we use the Lagrange equations! Using that ω^{body} formula, we write T and L = T-V directly in terms of the three Euler angles. Since the torque due to gravity is always along the line of nodes (**r** x **F** right hand), we know that L_ψ and L_φ (generalized momenta) are constants of the motion, see page 165, which corresponds to the fact that L is cyclic in φ and ψ. A formula for total energy E = T+V is also written and it is a third constant of the motion.

Various other constants now appear. First, we have ω-like a and b both in units of sec^{-1}:

$L_\psi = I_3\omega_z \equiv I_1 a$	defines a	// ω_z (body) is a constant for the top
$L_\varphi = I_1 b$	defines b	// Note that $a = (I_3/I_1)\omega_z$

we then quickly arrive at this fact

$\dot{\phi} = (b - aC_\theta)/S_\theta^2$	5-50	// eliminate $\dot{\psi}$ between 5-46 and 5-47

$\dot{\psi} = \omega_z - C\theta\,\dot{\phi}$ 5-51 // comes from ω^{body} form, see below

The idea would be that, if you can solve somehow for θ(t), you use 5-50 to get φ(t) and then 5-51 for ψ(t) and you have completely solved the problem. Notice from 5-50 that if θ(t) moves some relatively modest amount in its range, $\dot{\phi}$ generally has always the same sign, and this means the top is precessing.

G defines E' = E - ½ $I_3\omega_z^2$ = constant of motion. He then defines $\alpha = 2E'/I_1$ and $\beta = 2Mg\ell/I_1$ having dimensions sec^{-2} and we end up with a very simple equation for u = cosθ which is that $\dot{u}^2 = f(u)$ where f(u) is the cubic shown in 5-55 which involves all four of our constants a,b,α,β.

At this point, we get a very nice **qualitative discussion** of the solution u(t) = cosθ(t) of this equation. It can be done exactly I think with elliptical integrals and you then get your θ(t) and you can complete the program. But qualitative is better!

The idea is that f(u) must have the general shape shown page 167 where two of the three roots of f(u)=0 are physical and lie in the range (-1,1). Because $\dot{u}^2 = f(u)$, we know that u_1 and u_2 are turning points where $\dot{u}$ and hence $\dot{\theta}$ have zero "velocity" and change direction. Thus, presuming that we also will be getting some kind of precession, we know that the "figure axis" (spin axis of the top) will describe a **nutating path** as shown top of page 168. So already we have a basic understanding of the cause of nutation along with our precession for the top. If we arrange the constants a,b,α,β to make $u_1 = u_2$, then we will get precession without nutation.

Now let's digress back to our ω^{body} formula:

$$\omega_\phi^{body} = \dot{\phi}\,(\sin\psi\sin\theta,\ \cos\psi\sin\theta,\ \cos\theta) = \dot{\phi}\,\hat{n}_\phi$$
$$\omega_\theta^{body} = \dot{\theta}\,(\cos\psi,\ -\sin\psi,\ 0) = \dot{\theta}\,\hat{n}_\theta$$
$$\omega_\psi^{body} = \dot{\psi}\,(0,\ 0,\ 1) = \dot{\psi}\,\hat{n}_\psi$$

$$\omega^{body} = \dot{\phi}\,\hat{n}_\phi + \dot{\theta}\,\hat{n}_\theta + \dot{\psi}\,\hat{n}_\psi$$

where we are expressing things in "body coordinates". Notice that ω_z gets two contributions! The obvious one is just $\dot{\psi}$ from the last term, but another contribution $\cos\theta\,\dot{\phi}$ comes from the first term. Imagine a top that is not spinning at all, $\dot{\psi} = 0$, but is precessing due to some external contrivance. Imagine the tip down angle is only θ = 1 degree. The top basically has $\omega_z = \dot{\phi}$ in this case (cosθ = 1). At a general angle θ, $\dot{\phi}$ continues to make a contribution to ω_z. The instantaneous ω_z must be measured in an inertial frame momentarily aligned with the body while it is at some point its precession, and $\dot{\phi}$ always makes a contribution to ω_z. This is a tricky thing to understand.

In any event, we see from the above that $\omega_z = \dot{\psi} + C\theta\,\dot{\phi}$. We know that ω_z is a constant, so during the top's nutation, we are going to have "energy" shifting back and forth between $\dot{\psi}$ and $\dot{\phi}$ such that the sum ω_z stays constant. This is 5-51 quoted above. Usually $\dot{\psi}$ is large and does not change sign, but we might have $\dot{\phi}$ change sign during nutation as in the middle case page 168. The right case is what you get when you start a top by holding it and then "dropping" it.

<u>At this point G addresses some particular questions.</u>

1. **What happens when you drop-start a "heavy top"?** We know our upper angle θ_0 and we try to find the lower angle θ_1 with $u_1 = \cos\theta_1$. We get a messy quadratic for $x_1 = u_0 - u_1$ (the drop delta), but with the fast top assumption, we get the result x_1 shown in 5-62. The fast top means that the delta in angle θ will be small, so we can replace $\theta = \theta_0$ in certain places in the calculation of things. Continuing in this approximation, we then find a solution for $u(t) = \cos\theta_0 - (x_1/2)(1 - \cos(at))$ and we see that u(t) which is the nutation height is doing a simple trig motion with frequency none other than our constant a.

nutation frequency of heavy dropped top = $a = (I_3/I_1)\omega_z \sim \omega_z$

So faster top means faster nutation frequency. The nutation amplitude is

nutation amplitude of heavy dropped = $x_1/2 = (I_1/I_3)(Mg\ell/I_3\omega_z^2)\sin^2\theta_0 \sim 1/\omega_z^2$

so the faster top loses amplitude quickly.

Now that we in essence have found $\theta(t)$, we know $\dot{\phi} = (b - aC_\theta)/S_\theta^2$ and this tells us that $\dot{\phi} = (Mg\ell/I_3\omega_z)(1 - \cos(at))$ which does the same trig thing as u(t). Clearly then

average precession frequency of heavy dropped top = $(Mg\ell/I_3\omega_z) \sim 1/\omega_z$

In practice, small nutation gets damped out by friction and you have pseudoregular procession in this case.

2. **How might you launch a heavy top to get regular precession** (ie, no nutation at all). This corresponds to $u_1 = u_2$ or $\dot{\theta} = 0$. It turns out you have to sort of "throw" the top in one of two special ways, and you can then get no nutation.

3. **What happens if you start a top vertically and friction slows it down?** G shows that the answer is that the top stays vertical as long as $\omega_z > \omega_{critical}$ = shown in 5-71 page 174. As it slows down and there is loss of ω_z, it goes into nutating mode where $\theta_0 = 0$ and θ_1 keeps going lower and lower, the top wobbles and eventually falls over.

4. **G comments on three applications.** The **gyroscope** has fully free gimbals so no torque and you start a no-nutation motion and the thing points in the same way no matter what, The **gyrocompass** is a fancier thing in older designs. The last application is that the **sun and moon** put torques on the oblate earth and cause a very slow precession with a period of 26,000 years, but we don't actually do this problem. This is the **precession of the equinoxes** problem.

In section 5-8, G considers the motion of a **rotating electric charge distribution** in the presence of a uniform **magnetic field B.** A simple example is a current loop where the charges are electrons and the loop makes a moment M. Imagine maybe a superconducting loop with no battery needed. The loop of charge is the "rotating rigid body" here, and it will precess as the usual Larmor frequency. There is a small detail that it might have a very small nutation in some situation but we have to go read a 1951 G

paper to learn more about that. This is really an E&M topic, but it does sort of fit into this chapter. This book of course is just a set of lecture notes assembled into a reasonable structure and that is where Larmor precession ended up.

Chapter 6. Special Relativity in Classical Mechanics (p 185)

Goldstein uses the convention that $x_\mu = (\mathbf{x}, ict)$ which is a trick to get $x^2 - c^2t^2$ = world scalar so same in two different inertial frames. My later teachers always used $g_{\mu\nu}x^\mu x^\nu$ and all that stuff, but the results are all the same. Here are the topics addressed in this chapter:

- derivation of the 4x4 orthogonal transformation matrix for a z-boost, thus expressions for z' and t'.
- why this leads to Lorentz contraction of length, and time dilation of clock ticks, both when a rod or clock is viewed from a boosted frame.
- the rule for **velocity addition** for boosts in the same direction
- The notion of **proper time** τ which is a world scalar
- The notion that $u_\mu = dx_\mu/d\tau$ is a **4-velocity**, and $p_\mu = mu_\mu$ the **4-momentum**
- The idea that $F = ma$ becomes $d(mu_\nu)/d\tau = K_\nu$, **generalization of Newton's Law**
- For free particle, $T = \gamma mc^2$
- Statement of a **covariant Lagrangian** and **covariant E-L equations** to go with it,
  $L' = \frac{1}{2}\, mu_\nu u_\nu + (q/c)u_\nu A_\nu$ at least in the case of an EM potential
- rigid body motion has no obvious relativistic generalization.

Chapter 7. The Hamilton Equations of Motion (p 215)

In Section 7.1 we learn about the Legendre transformation as used to "change a coordinate" and this is written $g = f - ux$ for changing from x to u (and leaving other coordinates unchanged). Then we think of the definition $-H = L - p\dot{q}$ in exactly this way, where we want to change from $\dot{q}$ to p to get H(p,q). In this special case, when we do this and use the E-L equations and definition of p, we end up with the Hamilton equations of motion which are symmetric except for a minus sign,

$$\partial H/\partial q = -\,\dot{p} \qquad \partial H/\partial p = \dot{q} \qquad \partial H/\partial t = -\,\partial L/\partial t$$

In Section 7.2 we learn more about what this does for us. There are 2n truly independent variables the p and q, but there are 2n first order ODE's. In the L world, only the q's were really independent, and we then had n second order ODE's (the E-L equations). In the H world, cyclic coordinates are simpler. If some q does not appear in H, you can also eliminate the corresponding p since p = constant. In the L world, if some q did not appear in L, you still had its $\dot{q}$ to worry about. If you do $R \equiv H' = \Sigma^s\, \dot{q}p - L$ where the sum is only over the cyclic coordinates, you can then reduce your L problem to n-s coordinates, called Routh's Procedure, and R is the Routhian, one foot in each camp.

In Section 7.3 we compute some sample H's: (1) single particle in central force V(r), where the Ham equations reproduce our L-world results; (2) particle in an EM V. For non-rel, the V that you subtract

from T to get L has the form $V = q\varphi - (q/c)\mathbf{A}\bullet\mathbf{v}$ and so is a velocity-dependent potential. When this is recast into H(p,q), you find that $H = (\mathbf{p}\text{-}q\mathbf{A}/c)^2/2m + q\varphi$ and you see the "recipe" about how p is replaced with what is shown; (3) For rel particle, you again take the rel H and use the exact same recipe.

Finally, we start over using a covariant definition of H which we call $H' = p_\mu u_\mu - L'$ where L' was the covariant Lagrangian we used earlier. Everything seems to come out well here.

In Section 7.4 it is shown how you write $\delta(\int L dt) = 0$ in terms of H and instead of getting E-L equations, you get the Hamilton's equations.

In Section 7.5 Goldstein attempts to treat the **Principle of Least Action**. The "action" is $A = \int p\dot{q}dt = \int pdq$ and the Principle says that problem solution satisfies $\Delta(\int pdq) = 0$ when H is conserved, where $\mathbf{\Delta}$ is a new kind of variational operator different from δ which allows the time end-points to be varied as well as the path. Page 228 has a whole set of **interpretative comments** which I cannot justify. However, given the definition of the operator Δ, I am able to derive all equations, and he does prove that $\Delta(\int pdq) = 0$ when H is conserved, but I just don't understand what is really going on. He then shows how to recast the action's time integral as an integral over path length ds, so it is really a spatial thing, not a time thing. And he shows what the path length becomes for a multi-particle system, calls it $d\rho$.

On line at http://www.eftaylor.com/leastaction.html#variational there is a much better presentation and a good criticism of books like Goldstein which try to do it all formally with no insight as to what is really happening. So I will look at this Taylor website soon.

Chapter 8. Canonical Transformations (p 237)

8.1 A **canonical transformation** (CT, aka a **contact transformation**) from p,q,H to P,Q,K is one in which Hamilton's equations in the new coordinates with new Ham K are the same as in the old coordinates with H. This will be true if the modified Hamiltonian $p\dot{q} - H$ and $P\dot{Q} - K$ differ by dF/dt where F is called the **generator function** of the canonical transformation. Four cases i=1,2,3,4 are considered for functional dependence of F, and in each case we get three equations. The last equation in all cases is the same, that $K = H + \partial F_i/\partial t$. The first two equations for each F_i are simple relations of the form $A = \pm \partial F_i/\partial B$ where A and B are taken from the set p,q,P,Q, and these equations are different for the four cases. The case we will use the most is F_2 where we get $p_i = \partial F_2/\partial q_i$ and $Q_i = \partial F_2/\partial P_i$. In our cases of interest, F_i will not have explicit time dependence, so we will have K = H.

These CT relations are very similar in appearance to Hamilton's Equations $\dot{q}_i = \partial H/\partial p_i$ and $\dot{p}_i = -\partial H/\partial q_i$, but Hamilton's involve time derivatives on the coordinate side, and H instead of F_i. Notice that in the transformed system these will read $\dot{Q}_i = \partial K/\partial P_i$ and $\dot{P}_i = -\partial K/\partial Q_i$.

On page 243 one can interpret t and -H as the $n+1^{th}$ coordinate and momentum to get a compact form for the δ variation rule. It would be nice if in the P,Q,K new system you had lots of **cyclic coordinates**.

8.2 Examples. First is F = qP which implies the identity P=p etc. Second is f(q)P which gives a **point transformation**, an example of which is Q = Aq. Third is qQ which swaps q and p apart from a minus

sign, and shows how general they are. Fourth example is an obscure looking $F_1(q,Q)$ which results in the **harmonic oscillator's** H being $K = H = \omega P$, which is thus cyclic in Q so P is a constant of the motion = E/ω.

8.3 The **integral invariants** (invariant under a CT) are called J_1 through J_n. J_1 is shown page 247 and is some kind of integral over a region of phase space but there is a sum of dq dp in there. G proves in a lengthy math section that this J_1 is invariant by showing that the Jacobians shown in 8-34 are equal. He gives an ambiguous definition of J_2 which I shall ignore, but his definition of J_n is clearer: it is the volume of phase space within a given boundary surface. This volume is the same no matter which set of p and q you use.

8.4 Here the **Lagrange** {a,b} and **Poisson** [a,b] **brackets** are defined and they are computed for basic p,q arguments with the traditional very simple results which in QM would apply to the commutator [x,p] were we to set $i\hbar = 1$. When arguments are p's and q's, brackets are called "fundamental". If you treat the brackets as matrix elements for arbitrary but independent functions u_i and u_j , you find that LP = 1 so just as it looks, you can think of the Poisson bracket as the "inverse" of the Lagrange bracket. These brackets are all defined as differences of product pairs of partial derivatives.

For the Poisson bracket, you can put arbitrary F,G inside and [F,G] is exactly what you would expect, and is again invariant. If you take as G the p or q, you get things like $[F,p] = \partial F/\partial q$ if you just examine the four partial derivatives in the definition.

8.5 Using the Poisson definition and Hamilton's equations, we find things like:

$[q,H] = \dot{q}$

$[p,H] = \dot{p}$ and generally $du/dt = [u,H] + \partial u/\partial t$

so amazingly, the Poisson bracket with the Hamiltonian tells you the rate of change of a function u(p,q; t), but of course this is just "passing through" from the two fundamental results just stated. This section ends with a proof of **Poisson's theorem** which says if u and v are constants of the motion, then so is [u,v]. This makes use of **Jacobi's Identity**, trivial for commutators, but proven here for the Poisson bracket definition.

8.6 Here we consider an **infinitesimal CT** which results in δq and δp and we write the generator function $F_1(q,P) = \varepsilon\, G(q,p)$, where G stands for "Generator". We find that $\delta p = -\varepsilon\, \partial G/\partial q$ and similarly for δq with plus sign. In general, we discover that for some arbitrary function u(p,q) that $\delta u = \varepsilon[u,G]$. As a special case, this says $\delta H = \varepsilon[H,G]$ so if [H,G] = 0 in some problem H for some generator G, then we know that G must be a **symmetry** of the motion since δH = 0, ie, H = constant as you fiddle with G. For example, for a rot sym Hamiltonian, if G is a rotation generator, a rotation leaves H unchanged in value. That is, we end up with $H(P(p,q),Q(p,q)) \equiv H'(p,q) = H(p,q)$ = same value though arguments are different. Goldstein then does two examples: G = **linear momentum**, and then G = **angular momentum**. It all agrees with what we already know, just checks on the formalism really. $d\theta G = d\theta\, \mathbf{n}\bullet\mathbf{L}$ for rotation about **n** axis.

8.7 For quantum operators L_i we are used to various commutator relationships. In this section, all those usual things are derived for the Poisson brackets, and the results are the quantum results but with the quantum value $i\hbar = 1$.

8.8. A proof of **Liouville's Theorem** of statistical mechanics. First, you represent an ensemble of prepared systems as a cloud of dots in phase space which are close together. If you pick a particular point in phase space, you will find some density D of cloud points. The theorem says that as the system moves in time, the density D at any given phase space point does not change. The proof uses the page 267 picture. As time moves ahead, a boundary of systems evolves into a new boundary. But time evolution can be interpreted as a contact transformation (synonym for canonical transformation) with H as the generator function, and we know from the J_n Poincare invariant that the volume of phase space enclosed by a surface is invariant under such a transformation, although I admit the exact statement of this invariant's definition is a little hazy. Since the volume dV is therefore invariant, and since no systems leave or enter the boundary, dN is invariant, so D = dN/dV is also invariant.

Chapter 9. Hamilton-Jacobi Theory (p 273)

This is a very long chapter, and I have intentionally put some sections out of order in my review here to make things clearer to me.

9.1 The Hamilton-Jacobi Identity for the Hamilton's Principle Function S (273). Recall that the general canonical transformation does K = H +∂F/∂t. In the last chapter we pretty much ignored F's which were time dependent so we had K = H, but in *this* chapter we want to select F=S such that K = 0 [later in the chapter, we will look at F=W and K=E] . Our motivation is that, due to Hamilton's Equations on K, such an F would result in a CT where all the P_i and Q_i are constants, a fascinating idea in itself. If we look for an $F_2(q_i,P_i,t)$ that does this, we know that $p_i = \partial F/\partial q_i$, and if we rename F_2 to be $S(q_i,P_i,t)$, and if we call the constants $P_i = \alpha_i$, then K=0 is the same as saying $H(q_i, \partial S/\partial q_i) + \partial S(q_i,\alpha_i,t)\partial t = 0$. This *is* the **Hamilton-Jacobi equation** (**HJE**), and S is called **Hamilton's *Principle* Function** and $Q_i = \partial S/\partial \alpha_i \equiv \beta_i$. It turns out that dS/dt = L, the Lagrangian, and S = $\int$L dt = the **"true action"**. It will turn out that going to constants P and Q gives a powerful method to solve problems, and also has impressive implications for quantum theory. You will need to compute S somehow, but you won't be able to do that from S = $\int$L dt because you don't know the q(t) yet that you need inside L.

9.3 Hamilton's Characteristic Function W (279) Intentionally out of order. If our system of interest has constant energy E, we can write $\mathbf{S(q_i,\alpha_i,t) = W(q_i,\alpha_i) - Et}$ and then $\partial S(q_i,\alpha_i,t)/\partial t = -E$ and the HJE then says $H(q_i, \partial W/\partial q_i) = E = \alpha_1$ in G's notation. W is called **Hamilton's *Characteristic* Function**.

At this point, G starts talking about a *different* canonical transformation -- that generated by W, not S. For *this* transformation, we have $K = H + \partial W/\partial t = H = E = \alpha_1$. Recall for S we got K = 0. If we use W as the CT generator, we still get P_i = constants, still called α_i, including $P_1 = \alpha_1 = E$. Notice the use of i=1 to refer to this special constant.

We know that $\dot{P}_i = \partial K/\partial Q_i$ and $\dot{Q}_i = \partial K/\partial P_i = \partial K/\partial \alpha_i$. Since $K = H = \alpha_1$, we find right off the bat that $\dot{Q}_i = \delta_{i,1}$ so that $Q_i = \delta_{i,1}\, t + \beta_i$. Therefore, the W transformation takes us to a place *similar to* where S took us, but now one of the Q_i is non-constant and it is $Q_1 = t + \beta_1$ -- a simple linear motion. The other equations are $Q_i = \beta_i$. You can alternatively choose a set of constants $\gamma_i(\alpha_i)$ none of which is the energy α_1, and this causes K to have a slightly different functional form which we interpret as a slightly different canonical transformation, and we then end up with $P_i = \gamma_i$ and $Q_i = \nu_i t + \beta_i$ where ν_i =

$\partial\alpha_1/\partial\gamma_i$. Notice that all the Q_i in general might end up being linear, and now we have some frequencies ν_i. Think how much simpler this is than say some q(t) elliptic function or worse.

These two approaches, S and W, are summarized on pages 282-284 in a large **side-by-side table**. You can only use the W method when E = constant = α_1.

9.2 A HJE example: the harmonic oscillator (277) . Now that we have introduced the W method, here is an example of both the S and W methods. We know H, we write the HJE in S, write it again in W and we quickly have integrals for both S and W. At top page 278 we see $\beta_1 = \partial S/\partial\alpha_1$ = integral(q) – t and we solve this and get the expected harmonic oscillator q(t). Using the W method, we would have said $t + \beta_1 = \partial W/\partial\alpha_1$ = integral(q) and we would get exactly the same answer of course. $\omega\beta$ turns out to be the arbitrary phase angle $\cos\omega(t+\beta)$. He then concludes this section showing for this example that $S = \int L\, dt$ as claimed, something we can show only after the problem is solved for q(t).

9.4 Separation of Variables in the HJE (284) Suppose we could find $W = \Sigma_i W_i(q_i,P_i)$ in a problem where $H = H(q_1, p_i)$ -- that is, q_1 is the only non-cyclic coordinate. For other i, we know that $\partial W_i/\partial q_i = p_i = \alpha_i$, constants. Thus, $W_i = \alpha_i q_i$ is a candidate (no constant adder) W_i., and $W = W_1 + \Sigma_i\alpha_i q_i$. I think the only point of this little discussion so far is that any "cyclic" coordinates can be handled in a $\Sigma_i\alpha_i q_i$ part of a separable W. Our earlier example was to think $q_1 = t$ and $p_1 = -H = -E$ so this would say $W = W_1 – Et$. But that is exactly what we did when we said S = W – Et. So if we have a cyclic coordinate, we want to break out its little contribution to W in this manner.

Example: (p287) Central force problem with any potential V(r) with conserved E. $H(r,\varphi,p_r,p_\varphi)$ is cyclic in φ, so we break off its piece and write $W = W_1(r) + \alpha_\varphi\varphi$ where $\alpha_\varphi = p_\varphi = \boldsymbol{\ell}$ of our previous discussion of central forces. [We have in effect already broken off the -Et part going from S to W since H is cyclic in t] The HJE then involves only this W_1 as shown in 9-25b or 9-28 and we solve for W in which $E = \alpha_1$ appears. We then write our two little equations from page 281 that $Q_1 = t + \beta_1 = \partial W/\partial\alpha_1$ and $Q_\varphi = \beta_\varphi = \partial W/\partial\alpha_\varphi$. These equations have the form $t + \beta_1 = f(r,E,\boldsymbol{\ell})$ and $\beta_\varphi = g(r,E,\boldsymbol{\ell}) + \varphi$. The first is the central force solution giving r = r(t), the second gives $\varphi = \varphi(r)$ which is the orbit equation. These results are obtained very quickly here with minimal muss and fuss.

9.5 Action-angle variables (288). Many problems like Kepler or the 3D HO result in periodic motion in the various **phase-space planes $\mathbf{q_i}$-$\mathbf{p_i}$.** This can involve closed orbits called **libration**, or a repeated orbit like a sine wave called **rotation**; let's just consider the libration case now. Pendulum can do both. We also want to assume the notion that the coordinates of the problem are **"separable"** in the sense discussed earlier. That is to say, we assume we have $W(\mathbf{q},\mathbf{P},\boldsymbol{\alpha}) = \Sigma_i\, W_i(q_i, \boldsymbol{\alpha})$ and don't depend on **P** and **α** are just some assumed constants, ***and*** our problem can be separated into separate equations in each coordinate which G writes as $H_i(q_i, \partial W_i/\partial q_i, \boldsymbol{\alpha}) \equiv \alpha_i$, see my comments in confusion.doc.

We can then look at the CT equations. One says $p_i = \partial W_i(q_i,\boldsymbol{\alpha})/\partial q_i$. This has the form of the equation $p_i = f(q_i,\boldsymbol{\alpha})$ which, for a given choice of **α**, is the *equation of* some closed orbit in the q_i-p_i plane. [Notice that we can only have this if we have a *separable* W, otherwise you don't get private orbits in the private i-planes! That is, p_i would depend on all the q_j.] Note that p_i is not a constant.

At this point, we define some the **integrals** $J_i = \oint p_i dq_i$ which are integrated around exactly one cycle of the periodic motion in the i plane. By replacing $p_i = \partial W_i(q_i,\boldsymbol{\alpha})/\partial q_i$, we see that the J_i are all just

functions of the α_i, so we can write $J_i(\boldsymbol{\alpha})$. The main idea is they are constants. We would like to select our functions $W_i(q_i, \boldsymbol{\alpha})$ such that these constants J_i are the P_i of a CT. We can recast the form then and talk about $W_i(q_i, \mathbf{J})$ and this is our F_2 form with $\mathbf{P} = \mathbf{J}$. The other CT equation is $Q_i = \partial W/\partial P_i$ so we give Q the name "w" and write this as $w_i = \partial W/\partial J_i$. We started with some $H(\mathbf{p},\mathbf{q})$ and I guess we then end up with some $K(\mathbf{J},\mathbf{w})$. It seems that $\partial W/\partial t = 0$, so $K(\mathbf{J},\mathbf{w}) = H(\mathbf{p},\mathbf{q})$, and in another document I have discussed "what this means" in the sense of K = H. However, we know that $\dot{J}_i = \partial K/\partial w_i = 0$ since the J_i are constant, and therefore K must be cyclic in the w_i, so we can write $K(\mathbf{J},\mathbf{w}) = K(\mathbf{J})$. So we might say H = K(J) and an example of this appears on page 303 equation 9-70, more generally 9-36 page 291. Since we are doing W-theory, we are assuming that $H = E = \alpha_1$.

We know that $\dot{w}_i = \partial K(\mathbf{J})/\partial J_i = f(\mathbf{J})$ which we call $\nu_i(\mathbf{J})$. Then of course we have the desired result that says $w_i = \nu_i(\mathbf{J})\ t + \beta_i$. We can examine, just for our interest, the amount by which w_i changes during one libration cycle and write $\Delta w_i \equiv w_i(t+\tau) - w(t) = \nu_i(J_i)\ \tau$. By the marvelously simple argument given on page 292 bottom, we can show that over one cycle in the p_iq_i plane, $\Delta w_i = 1$ and we therefore find that $\nu_i(J_i)\ \tau = 1$ so we can interpret ν_i as the frequency of the repeated motion in the q_ip_i plane. This is the big payoff of action angle variables.

Since $J_i = \oint p_i dq_i = \oint p_i \dot{q}_i dt$, we see that J_i can be interpreted as a component of the total abbreviated action S_0 which is associated with coordinate q_i. Thus, $\Sigma_i J_i$ is the total action S_0 again time-integrated around the one-cycle path. Thus, the name "action variable" for J_i makes total sense. It has the dimensions of angular momentum, so the conjugate variable w_i is thus an "angle". Note that the J_i are not angular momentum **r** x **p** in any sense, but action does have the *dimensions* of angular momentum. You cannot say $L_i = f(p_i, \dot{q}_i)$ because L mixes coordinates together, $L_z = xp_y - yp_x$.

9.7 The Kepler problem in action-angle variables (299) Out of order intentionally. We have just learned a giant glob of theory, let's not wait to apply it to Kepler 3D. As noted in my "clarification of separation" in confusion.doc. we are going to write

$$W(r,\theta,\varphi, \boldsymbol{\alpha}) = W_r(r,\boldsymbol{\alpha}) + W_\theta(\theta,\boldsymbol{\alpha}) + W_\varphi(\varphi,\boldsymbol{\alpha})$$

where $\boldsymbol{\alpha}$ are some constants. We then come up with the three separated equations which G would write as

$H_\varphi(\varphi, \partial W_\varphi/\partial\varphi, \boldsymbol{\alpha}) = \alpha_\varphi;$	// see 9-63a	which involves only α_φ, one of our $\boldsymbol{\alpha}$
$H_\theta(\theta, \partial W_\theta/\partial\theta, \boldsymbol{\alpha}) = \alpha_\theta;$	// see 9-63b	which involves α_φ and α_θ
$H_r(r, \partial W_r/\partial r, \boldsymbol{\alpha}) = \alpha_1;$	// see 9-63c	which involves α_θ and α_1

and which I can interpret this way [the H's are not all Hamiltonians, and $H \neq \Sigma_i H_i$]

H_φ	$\hat{L}_z\ (p_\varphi = \partial W_\varphi/\partial\varphi) = \alpha_\varphi = m$
H_θ	$\hat{L}^2\ (p_\varphi = m,\ p_\theta = \partial W_\theta/\partial\theta) = \alpha_\theta^2 = L^2$
H_r	$H\ (p_r = \partial W_r/\partial r, L^2) = E$

so each equation is associated with a conserved quantity. This is all just set-up work.

Next, we compute the three J_i for this problem. Various tricks are used to do this, but we get in the end these results:

$J_\varphi = 2\pi\alpha_\varphi$
$J_\theta = 2\pi(\alpha_\theta - \alpha_\varphi)$
$J_r = -(J_\theta + J_\varphi) + \pi k\sqrt{2m/-E}$

where the two α_i constants are interpreted as shown above. Solving the last for E we get

$$E = K(J) = -2\pi^2 mk^2/(J_r + J_\theta + J_\varphi)^2 \qquad 9\text{-}70$$

Since $\partial K(\mathbf{J})/\partial J_i = \nu_i(\mathbf{J})$, we see that our three ν_i are degenerate, so call them all just ν. We do the derivative and sub in from above to find that $\tau = \pi k\sqrt{m/-2E^3}$. In the ellipse orbits, it turns out that the semi-major axis is A = k/2(-E), so we get 3-54 page 80 which is Kepler's Third Law. Obviously this action-angle stuff is going to give this result since we are computing libration period / frequency.

At this point, starting on page 304, G applies some extra "theory" presented earlier for handling degenerate ν_i. I will comment on that in the next section.

9.6 Further properties of action-angle variables (294) If an i-plane periodic motion has ν_i, we know it can only consist of this fundamental and harmonic multiples of ν_i, so you can do a Fourier series expansion for $q_i(t)$ if you want. Back in Cartesian space, some 2D problem $x_i(t)$ will be given by a double sum $\Sigma_{m,n}$. If you try to make this be periodic with some T, you find that ν_1/ν_2 has to be a ratio of integers, otherwise the resulting function is not periodic. This is the famous situation of being **conditionally** or **multiply periodic**. G uses the word "**incommensurable**" to mean the ratio of two numbers is not a rational number. This is an antique term meaning two numbers have an irrational ratio.

G then considers a situation in which you have n frequencies ν_i, but m of them are the same. In this case, you can do a little extra CT like this

q_i	p_i	→	Q_i	P_i
w_i	J_i		w'_i	J'_i

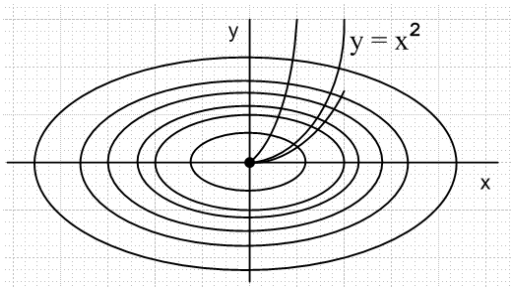
This maps the m degenerate frequencies (really the m duplicate ones) to $\nu_i' = 0$. I have done all the details here in a nice matrix notation, but the method does not seem crucial for our idea flow here. In the case of the Kepler motion, G shows this little extra CT on page 304. He computes the three w_i' variables and the three J_i' as shown, and the new K' Hamiltonian (he calls H as usual) is a function only of J_3' so that two of your ν_i' come out zero as intended. More on this in next section.

9.7 The Kepler problem continued comments. Page 305 now. G shows us that the w_1', w_2' and ratio J_1'/J_2' are in fact some fixed Euler angles we might call φ_0, ψ_0 and $\cos\theta_0$. The action-angle approach for astronomy is used even nowadays because it is well suited to perturbation analysis, the **Delaunay elements**. After the **Bohr atom** (quantized orbits with classical radiation, no photons) came out in 1913, he and **Sommerfeld** tried to improve the model by saying orbits were elliptical and the action J_i was quantized. As the atom started to break the triple degeneracy in the Kepler problem we just saw, you got to add new action quantization rules for each new J_i' (really $\nu_i' \neq 0$) that appeared. You could produce

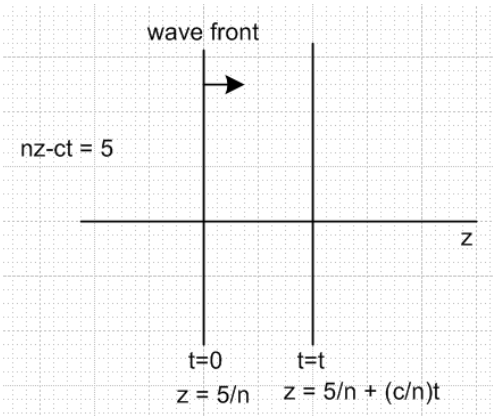
the hydrogen atom quantum numbers in this obscure way, but alas, when the real QM came out for solving the H atom (Pauli in 1925 by matrix, Schrodinger in 1926 by orbitals), all this action-angle stuff was shoved aside. So action-angle nowadays lives only in text books and in astronomy work.

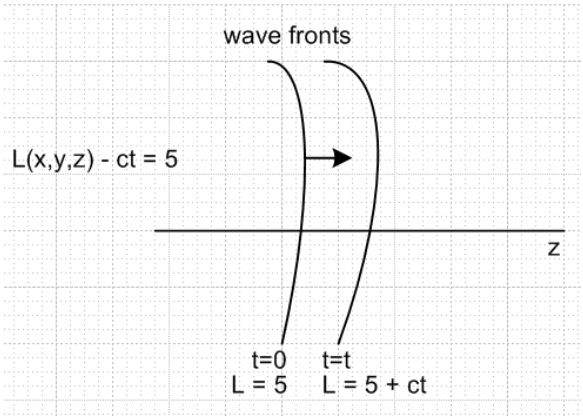
9.8 Hamilton-Jacobi theory, geometric optics, and wave mechanics (307)

In my notes, I first reviewed the whole notion of **"surfaces" and normals** to surfaces and wave fronts and such things. I learned how to compute the "normal curves" which are perp to a family of surfaces and I found that, for $f(x,y) = ax^2 + by^2 = z$, the normal curves are $h(x) = Kx^{b/a}$. Here, we generate a family of concentric ellipses as we let z take different values.



Next, I followed G's ***review of geometric optics***. You have the exact E&M wave equation for function φ, you consider a slowly varying index n(x) relative to λ, and you use a trial form for wave solution φ in which the phase is represented by L(**r**) replacing $n\hat{\mathbf{k}}\bullet\mathbf{r}$. In this case, if you consider a "constant phase" surface 5 = P(**r**,t) = L(**r**)-ct, you find that as t increases, the surface P=5 assumes a sequence of positions (and shapes) L = 5 + ct. I found this picture very helpful, comparing n= constant to slowly varying n:

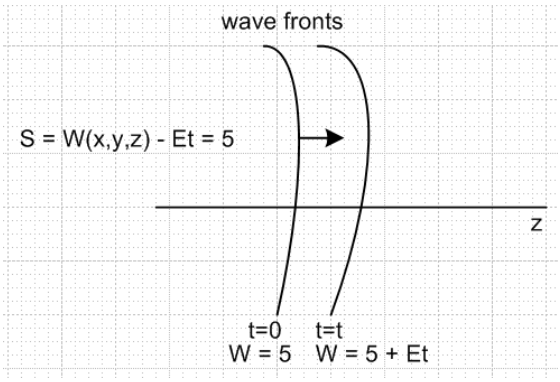


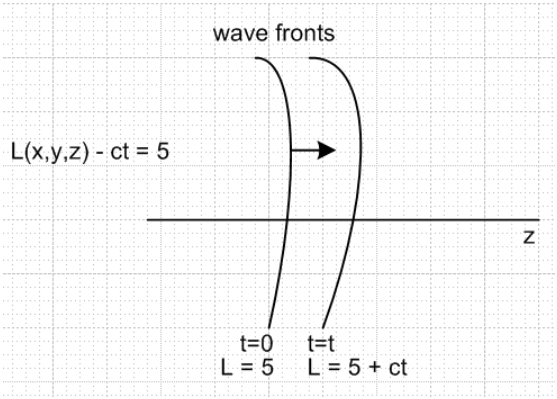


I showed that the normal to a wavefront is given by $\mathbf{n} = \nabla_{3D}L(\mathbf{r})/|\nabla_{3D}L(\mathbf{r})|$, and if we march a distance ds in the direction of this normal, we find that $dP(\mathbf{r},t) = |\nabla_{3D}L(\mathbf{r})|\,ds$, and finally we learn that the wavefront velocity at a point is given by $u(\mathbf{r}) = ds/dt = [dP/|\nabla_{3D}L(\mathbf{r})|] / dt = c/|\nabla_{3D}L(\mathbf{r})|$. Finally, in the approximation that n varies slowly relative to λ, we show that the surfaces L(r) must obey the equation $(\nabla_{3d}L(\mathbf{r}))^2 = n^2(\mathbf{r})$. This is called the **eikonal equation**, and L the **eikonal**, since L is the image phase in geometric optics, and eikon = icon = image.

Next, G connects all this stuff to the Hamilton-Jacobi theory. The HJE says $H(q_i, p_i = \partial W/\partial q_i) = E$ as we saw above. For a system consisting of a single particle mass m in a potential V(r), this HJE

equation becomes $(\nabla W(\mathbf{r}))^2 = 2m(E-V(\mathbf{r}))$ which looks just like $(\nabla_{3d} L(\mathbf{r}))^2 = n^2(\mathbf{r})$ of geometric optics. We then make these connections:



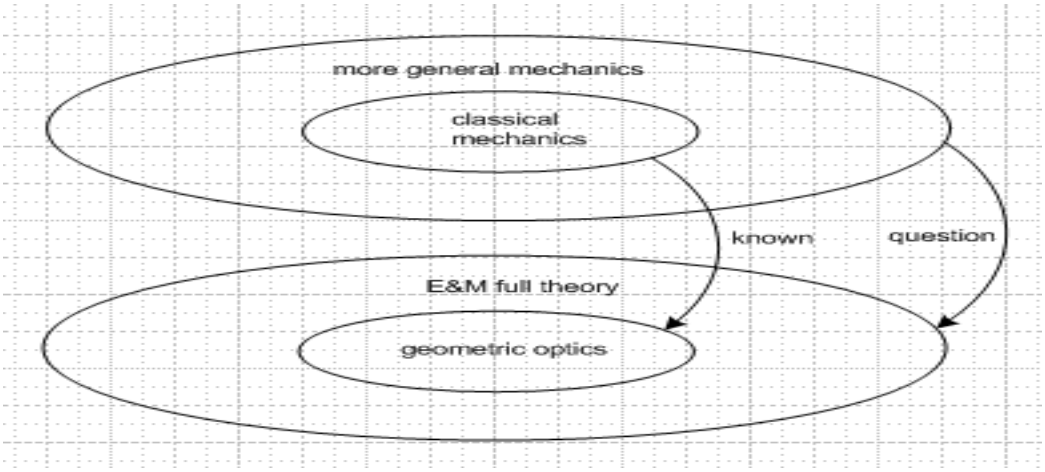


W-world	optics world
$S(\mathbf{r},t) = W(\mathbf{r}) - Et$	$P(\mathbf{r},t)$ = the phase = $L(\mathbf{r}) - ct$
$W(\mathbf{r})$	$L(\mathbf{r})$
E	c
$f = 2m(E-V(\mathbf{r})) = 2mT$	$f = n^2(\mathbf{r})$
$(\nabla W(\mathbf{r}))^2 = 2m(E-V(\mathbf{r})) = 2mT$	$(\nabla L(\mathbf{r}))^2 = n^2(\mathbf{r})$

Via this long theoretical path, we finally understand Goldstein's opening picture on page 308. The surface $S(\mathbf{r},t) = 5$ (he calls it a) is going to move to the right as t increases, taking the position of the various W = 5 + Et. So somehow *associated with* our little mass particle in a potential is a function S(r,t) which has wavefronts like S(r,t)=5 which move in time and take on the shapes of the surfaces W = a + Et. So ***entirely within the realm of classical mechanics, we are seeing some kind of "wave" associated with a particle.*** I never knew this! A simple calculation shows that the wave velocity is u = E/p where p = the usual mv. A particle of constant energy E then has a faster wave velocity as v slows down. We learn that $\mathbf{p} = \nabla W$ so the physical particle path must be along one of the normal curves.

Somehow, we suddenly have some kind of "wave theory" of the motion of a classical particle. At the same time, we have some kind of "particle trajectory theory" for light. And this has nothing to do with quantum mechanics. Another connection is the least-action business. In classical mechanics we already showed that the solution trajectory of a particle minimizes $\int \sqrt{2mT}\, ds$, and by our analogy above, this corresponds to the geometric optics rule called Fermat's Principle where we minimize $\int n ds$.

Big Question. This question could have occurred to people of the late 1800's who knew all the stuff described above, but who never heard of quantum mechanics. We have seen that "classical mechanics", as embodied in the Hamilton-Jacobi formalism, can be fully described by a function W which seems to define some kind of wavefronts which move in time, and we have seen that these waves satisfy a differential equation which corresponds to the "geometric limit" of E&M, it its optics incarnation. We have seen this particle-wave duality in classical mechanics, but the particles are the "senior partner" since nobody knows what to "do with" those waves in classical mechanics. The question is this: is there some more *general* theory of mechanics which corresponds to the full E&M theory?



Goldstein then proceeds to do what the 1800's people might have done, if they had motivation to do so, which they did not because classical mechanics was thought to be the full theory and no experiments at that time showed it to be incomplete.

1) call the waves that are associated with particles ψ, and we know from the classical theory that $\psi(\mathbf{r},t) = \exp[\, i\, S(\mathbf{r},t)/\hbar]$ where $\hbar$ could be any constant value you want. We expect these waves to satisfy the wave equation.

2) Compare these waves with those of geometric optics, the comparison being this:

$\psi(\mathbf{r},t) = \exp\{\, i\, 2\pi\, [\, W(\mathbf{r})/h - (E/h)t]\, \}$ // classical mechanics waves
$\varphi(\mathbf{r},t) = \exp\{\, i\, 2\pi\, [\, L(\mathbf{r})/\lambda_0 - \nu\, t]\, \}$ // geometric optics waves

3) observe that both W and L satisfy an eikonal equation. In classical mechanics, this eikonal equation is in fact the Hamilton-Jacobi equation for W for a single particle in a potential V.

4) Make the postulate that the frequency of the classical waves is not arbitrary, but is fixed by the relationship $\nu = E/h$ where h is some constant of the universe. This proportionality E to ν is strongly suggested by the above analogy. From this we get $\lambda = h/p$ which is the deBroglie wavelength of a particle -- this is the wavelength of those "associated waves". If you are working with large distances compared to λ, then classical mechanics is correct because it corresponds to the geometric limit of E&M. Otherwise, you have to worry about the "waves" that were the junior partners. You can have the same kind of things that occur in non-geometric optics, such as interference. [Evidence: Davisson-Germer 1927:]

In 1924 Louis de Broglie presented his thesis concerning the wave-particle, proposing the idea that all matter displayed the wave-particle duality of photons.[1] According to de Broglie, for all matter *and* for radiation alike, the energy E of the particle was related to the frequency of its associated wave ν, by the Planck relation

$$E = h\nu,$$

and that the momentum of the particle p was related to its wavelength λ by what is now know as the de Broglie relation

$$p = \frac{h}{\lambda},$$

where h is Planck's constant.

In 1926, Walter Elsasser remarked that the wave-like nature of matter might be investigated by electron scattering experiments on crystalline solids, in the same the wave-like nature of was confirmed through X-ray scattering experiments on crystalline solids.[1]

In 1927 at Bell Labs, Clinton Davisson and Lester Germer fired slow moving electrons at a crystalline nickel target.[2] The angular dependence of the reflected electron intensity was measured, and was determined to have the same diffraction pattern as those predicted by Bragg for X-rays. This was also replicated by Joseph J. Thomson. [1]

5) use this fact with other small results obtained from the classical mechanics theory to conclude that $\psi(\mathbf{r},t) = \exp[\, i\, S(\mathbf{r},t)/\hbar]$ satisfies the *Schrodinger equation*! This is pretty amazing really. [1926]

6) Going the other way, conclude that a photon is just a particle going the speed of light, and it also has the relationship $\nu = E/h$ and it travels on the normal curves of the optical wavefronts. [Evidence: photoelectric effect, resolved by Einstein in 1905:]

When a metallic surface is exposed to electromagnetic radiation above a certain threshold frequency, the light is absorbed and electrons are emitted. In 1902, Philipp Eduard Anton von Lenard observed that the energy of the emitted electrons increased with the frequency, or colour, of the light. This was at odds with James Clerk Maxwell's wave theory of light, which predicted that the energy would be proportional to the intensity of the radiation. In 1905, Einstein solved this paradox by describing light as composed of discrete quanta, now called photons, rather than continuous waves. Based upon Max Planck's theory of black-body radiation, Einstein theorized that the energy in each quantum of light was equal to the frequency multiplied by a constant, later called Planck's constant. A photon above a threshold frequency has the required energy to eject a single electron, creating the observed effect. This discovery led to the quantum revolution in physics and earned Einstein the Nobel Prize in 1921.

A later piece of evidence here was the Compton Effect, showing that a photon scatters off an electron and can have its energy changed, as if it were a particle.

The Compton effect was observed by Arthur Holly Compton in 1923 and further verified by his graduate student Y. H. Woo in the years following. Arthur Compton earned the 1927 Nobel Prize in Physics for the discovery.

Here is a collection of "facts" associated with the above discussion and derived in the raw notes:

$\sqrt{2mT} = p = mv.$	$\mathbf{p} = \nabla W$	momentum
$u = E/\lvert\nabla W\rvert = E/\sqrt{2mT} = E/p.$		wave phase velocity
$\lambda = u/\nu = E/(p\nu) = E/p)/(E/h) = h/p$		wavelength with the $E = h\nu$ assumption
$n = c/u = c/[E/\lvert\nabla W\rvert] = \lvert\nabla W\rvert\, c/\, E = pc/E$		index

Conclusion: This is a pretty profound section of Goldstein's book, showing that classical mechanics is the same limit of quantum mechanics that geometric optics is of optics. In order to even *understand this* you have to learn about Lagrangian mechanics and the Lagrange equations (Chap 1), and then about Hamiltonian mechanics and Hamilton's Equations (Chap 7), and then about canonical transformations which preserve Hamilton's equations (Chap 8), and then about the particular one which makes K = 0, called S and is identified with the action and called Hamilton's principle function, and the one W from S =

W-Et which takes you instead to K=E and is called Hamilton's characteristic function (Chap 9). The whole thing is strengthened when you see that geometric optics has its eikonal equation and this is really the same as the Hamilton-Jacobi equation for a particle in a potential V. Before reading this stuff, I had no idea that there were some waves associated with classical particles within classical mechanics!

Chapter 10. Small Oscillations (p 318)

10.1 Formulation of the Problem (318) We are in some arbitrary generalized coordinates q_i, we define some offset coordinates $\eta_i = q_i - q_{i0}$ around equilibrium values at which we set $V(\mathbf{q}_o)=0$. We have some arbitrary conservative potential V and some arbitrary kinetic energy T. We define symmetric matrices

$$T_{ij} \equiv \partial_i \partial_j V(\mathbf{q}_0) \qquad M_{ij} = \tfrac{1}{2} \Sigma_k m_k \, \partial \mathbf{r}_k/\partial q_i \bullet \partial \mathbf{r}_k/\partial q_j \,|^{q=q_o} = M_{ij}$$

Taylor expand to find that $V = V \approx \tfrac{1}{2} \boldsymbol{\eta}^T V \boldsymbol{\eta}$ and, if scleronomous coordinate equations, have $T = \tfrac{1}{2} \dot{\mathbf{q}}^T M \dot{\mathbf{q}}$, so $L = T\text{-}V = \tfrac{1}{2} \dot{\boldsymbol{\eta}}^T T \dot{\boldsymbol{\eta}} - \tfrac{1}{2} \boldsymbol{\eta}^T V \boldsymbol{\eta}$ and EL says $T\ddot{\boldsymbol{\eta}} + V\boldsymbol{\eta} = \mathbf{0}$. We end up with a system of ODE's in the η_i.

10.2 The eigenvalue equation and the principle axis transformation(321) We try $\boldsymbol{\eta}(t) = C \, \mathbf{a} \, e^{-i\omega t}$ which results in $(V - \omega^2 T) \, \mathbf{a} = \mathbf{0}$ or $(V - \omega^2 T) \, \boldsymbol{\eta} = \mathbf{0}$. Then $\det(V - \omega^2 T) = 0$ determines the eigenvalues ω_k^2 and we have $(V - \omega_k^2 T) \, \mathbf{a}_k = \mathbf{0}$, a vector equation for each k. Given V and T, we can find the eigenvectors from $(T^{-1}V - \omega_k^2) \, \mathbf{a}_k = \mathbf{0}$ using Maple. We can form matrix A from $A_{jk} = (a_k)_j$. If the eigenvectors are unit normalized, then A is a rotation that brings the matrix $T^{-1}V$ to diagonal form, but this is not what we want to do. We will use a different normalization!

We first write $VA = TA\Lambda$ which is the matrix form of our eigenvalue equation, where Λ is a diagonal matrices whose elements are the ω_k^2. We then digress to show that $B \equiv A^\dagger TA$ is a real diagonal matrix. Showing it is *real* is trivial given that T is symmetric, as shown in Lemma 1. In Lemma 2 we show that the ω_k^2 *are real* (though not necessarily positive) so Λ is real. Then from Lemma 3 we show that $A^\dagger TA$ is *diagonal* because it commutes with the diagonal matrix Λ.

Once we know that $B \equiv A^\dagger TA$ is real and diagonal (and positive definite, meaning diags all > 0), we can define a new transformation $A' = AS$ where S is a diagonal stretching matrix $S_{ij} = \delta_{ij}/\sqrt{B_{ii}}$. This amounts to selecting a special normalization for the eigenvectors $\mathbf{a}_k'$ and we get $B' \equiv A'^\dagger TA' = 1$. We no longer have any freedom in selecting normalization of the $\mathbf{a}_k'$, we have used that freedom up now.

From now on, we drop the prime and assume that A refers to this A'.

The payoff is seen going back to matrix equation $VA = TA\Lambda$ which is true in any normalization of the $\mathbf{a}_k$, but here we imply our special one. If we apply $A^\dagger$ on the left, we get $A^\dagger VA = \Lambda$ so our special matrix A diagonalizes *both* T and V, sending $T \rightarrow 1$ and $V \rightarrow \Lambda$. This special matrix A is a **principle axis transformation** since it diagonalizes our problem.

10.3 Frequencies of free vibration and normal coordinates (329) Go back to $(V - \omega_k^2 T) \, \boldsymbol{\eta} = \mathbf{0}$ and replace $\boldsymbol{\eta} = A \, \boldsymbol{\zeta}$ where $\boldsymbol{\zeta}$ are new, transformed coordinates obtained from $\boldsymbol{\zeta} \equiv A^{-1}\boldsymbol{\eta}$. Apply $A^\dagger$ to the left and we get $(\Lambda - \omega_k^2 1) \, \boldsymbol{\zeta} = 0$ which is a fully diagonalized equation. We can look back at our trial solution which was $\boldsymbol{\eta}(t) = C_k \, \mathbf{a}_k \, e^{-i\omega_k t}$ and we find that $\boldsymbol{\zeta} \equiv A^{-1}\boldsymbol{\eta}$ which leads to

$\zeta_i = C_k \exp(-i\omega_k t)$ $\quad i = k$

$\zeta_i = 0$ $\quad i \neq k$

Thus, in the kth **"normal mode"**, all components of the ζ vector are 0 except component k which is doing harmonic oscillator at ω_k (and no harmonics). The coordinates $\zeta \equiv A^{-1}\eta$ are the **"normal coordinates"** and we get $T = \frac{1}{2} \Sigma_i \dot{\zeta}_i^2$, $V = \frac{1}{2} \Sigma_i \omega_i^2 \zeta_i^2$ and $L = \frac{1}{2} \Sigma_i (\dot{\zeta}_i^2 - \omega_i^2 \zeta_i^2)$, so the EL equations are now $\ddot{\zeta}_i + \omega_i^2 \zeta_i = 0$, confirming our harmonic oscillator. The ODE's are fully decoupled by the diagonalization process. Compare to our messy starting problem of $\Sigma_j (T_{ij} \ddot{\eta}_j + V_{ij} \eta_j) = 0$.

We then show that if we are at a V minimum, the ω_k^2 are positive, so we get HO. If at a max, we get negative and get expo growth which is the instability.

This method is similar to but not the same as the inertia tensor stuff we did. For inertia, we called our transformation A by the name R, since it was a pure rotation, which put the ellipsoid $I = \mathbf{n} \bullet \mathbf{I} \bullet \mathbf{n}$ into alignment with the new axes so $I = I = \mathbf{n}' \bullet \mathbf{I}_{diag} \bullet \mathbf{n}'$. In this chapter, we do a similar thing with a rotation A, to get ellipse aligned in $\dot{\boldsymbol{\eta}}$ space, but then our stretch S takes it into a unit sphere so

$\dot{\boldsymbol{\eta}}' \bullet T_{diag} \bullet \dot{\boldsymbol{\eta}}' = \dot{\boldsymbol{\zeta}} \bullet 1 \bullet \dot{\boldsymbol{\zeta}}$ and this diagonalizes our entire problem.

10.4 Free vibrations of a linear triatomic molecule (333). We apply the theory of the last section to this little 3x3 problem (which is also treated in my M&M book). It is quick work to have Maple figure out the eigenvectors and eigenvalues of the problem $(T^{-1}V - \omega_k^2)\, \mathbf{a}_k$.

```
eigenvalues(Q);
```

$$\frac{k}{m}, 0, \frac{k(2m+M)}{Mm}$$

```
e :=eigenvectors(Q);
```

$$e := \left[\frac{k}{m}, 1, \{[-1, 0, 1]\}\right], [0, 1, \{[1, 1, 1]\}], \left[\frac{k(2m+M)}{Mm}, 1, \left\{\left[1, -2\frac{m}{M}, 1\right]\right\}\right]$$

At this point the eigenvalues have some offbeat normalization. We then find the stretching S matrix that gets us so $A^{\dagger} T A = 1$ and $A^{\dagger} V A = \Lambda$. In the end we are able to write down our specially normalized matrix A (see the triatomic.mws file)

```
A := simplify(evalm(A&*&*()));
```

$$A := \begin{bmatrix} -\frac{1}{2}\frac{\sqrt{2}}{\sqrt{m}} & \frac{1}{2}\frac{\sqrt{2}\sqrt{m(2m+M)M}}{m(2m+M)} & \frac{1}{\sqrt{2m+M}} \\ 0 & -\frac{\sqrt{2}\sqrt{m(2m+M)M}}{(2m+M)M} & \frac{1}{\sqrt{2m+M}} \\ \frac{1}{2}\frac{\sqrt{2}}{\sqrt{m}} & \frac{1}{2}\frac{\sqrt{2}\sqrt{m(2m+M)M}}{m(2m+M)} & \frac{1}{\sqrt{2m+M}} \end{bmatrix}$$

and $\boldsymbol{\eta} = A\,\boldsymbol{\zeta}$ are the normal motions, so $\eta = 1/\sqrt{2m}\,(-1,0,1)e^{-i\omega_1 t}$ for example for the $\omega_1 = k/m$ motion. This comes from the first column of A shown above. The nature of the other two motions are shown most simply in the eigenvectors above. Remember that our norm'd solutions are the columns of A.

We only solved the "trilinear" problem, but there are two equivalent sideways modes that we did not compute. G just wants to make sure we know they are there.

10.5 Forced vibrations and dissipative forces (338) We now step up to the next level of complication, adding "linear in velocity" frictional forces F,

$$T\ddot{\boldsymbol{\eta}} + F\dot{\boldsymbol{\eta}} + V\boldsymbol{\eta} = \mathbf{0} \qquad\qquad (V - i\omega_k F - \omega_k^2 T)\,\mathbf{a}_k = 0$$

Part 1: Suppose the A matrix *of the previous problem* diagonalizes T (to 1) V *and* F. If we install the *same* $\boldsymbol{\eta} = A\,\boldsymbol{\zeta}$ above top left we get

$$\ddot{\zeta}_i + F_i\dot{\zeta}_i + \omega_k^2\,\zeta_i = 0$$

where F_i are the diagonal elements of the diagonalized F. We are still completely diagonalized and we know how to solve this kind of equation by trying out $\zeta(t) = C\,\mathbf{a}\,e^{-i\omega t}$. This gives an expression for eigenvalues for ω which we might call ω_k' since not same as ω_k. The result is 19-68 for ω_k' which shows a shifted frequency and a decay factor as well

$$\omega_k' = \sqrt{\omega_k^2 - (F_k/2)^2} - i\,(F_k/2)$$

This means that oscillation in each normal mode will decay according to an F_k time constant. So we have in this case a complete normal modes solution to the problem. The modes don't cross couple.

Part 2: In the more likely case that our former A does *not* diagonalize F, the problem is too hard to solve in this book, which I guess is why G did not attempt it. I took a quick shot at it in the raw notes, using the "new A" of this new problem, but I got nowhere (see below). Staying with the old A, we get a set of cross-coupled ODE's of the following form:

$$\ddot{\zeta}_i + \Sigma_j F_{ij}\dot{\zeta}_j + \omega_k^2\,\zeta_i = 0 \qquad \text{or} \qquad \ddot{\boldsymbol{\zeta}} + F\dot{\boldsymbol{\zeta}} + \omega_k^2\boldsymbol{\zeta} = \mathbf{0}$$

The solutions will link the normal modes together, causing motion to move between them, much as is true for the original coordinates $\boldsymbol{\eta}$. If you install the trial solution $\boldsymbol{\eta} = A\,\boldsymbol{\zeta}$ and $\zeta(t) = C\,\mathbf{a}\,e^{-i\omega t}$ and write the result as a matrix equation, you get $TA\Omega^2 + iFA\Omega - VA = 0$ where of course A consists of the $\mathbf{a}_k$ eigenvectors of this new problem. Compare to $-T\ddot{\boldsymbol{\eta}} - F\dot{\boldsymbol{\eta}} - V\boldsymbol{\eta} = \mathbf{0}$ where eact ∂_t creates a $-i\omega$. The matrix Ω is diagonal with ω_k elements. But I don't know what to do with this matrix equation. It seems that we can normalize A still so that $B = A^\dagger TA = 1$ so you get $\Omega^2 + i\,A^\dagger FA\Omega - A^\dagger VA = 0$. but then I am dead in the water. I know there are good methods to deal with this stuff, but now is not the time for such a diversion.

Goldstein then considers the general case (Part 2) where you apply a **driving force**. In this case, everything runs at the driving frequency and you can solve the problem using Cramer's Rule and see the various damped resonances. Compare the two problems, hard and easy:

$$(V - i\omega_k F - \omega_k^2 T)\,\mathbf{a}_k = 0 \qquad\qquad (V - i\omega F - \omega^2 T)\,\mathbf{a} = \mathbf{f}(\omega)$$

where $\mathbf{f}(\omega) = \mathbf{C}\, e^{-i\omega t}$ is some driving force running at ω. The problem on the right is Cramer's rule, the problem on the left (homogeneous Fx = 0) is an eigenvalue problem that we have no simple way to solve, but we do have brute force methods I am sure, perhaps Stakgold has the answers.

G's final comment is that this type of analysis is being used in the EE world, but he gives no examples, saying to see other books such as his second last reference.

Chapter 11. Continuous Systems and Fields (p 347)

11.1. Transition from discrete to continuous system (347). G deduces the Lagrangian for an elastic rod for longitudinal vibrations. The number of mass points x_i spaced by Δx = 1 goes to infinity, so the number of generalized coordinates $\eta_i = \eta(x_i)$ becomes infinite, called η(x), sum becomes integral and integrand of L is called $\mathcal{L}$, m/a → μ, linear mass density, ka → Y, the Young modulus. In similar fashion, G takes the limit of the EL equation with finite Δx and gets the 1D wave equation, and we learn the velocity in a rod.

11.2 The Lagrangian formulation for continuous systems (350) The general $\mathcal{L}$ is considered a function of η, $\dot{\eta}$, **r**, t and also of ∇η which is *something new* and arises from deltas between adjacent η coordinates when you take the limit. The integral is $d^3\mathbf{r}$ to get L, does not include time t. However, if you add dt you get the action, and then you do a variational δI=0 on this and you obtain the Hamilton's Principle equations which have the form $\partial_t(\partial_{\dot{\eta}}\mathcal{L}) - [\,\partial_\eta\mathcal{L} - (\nabla\bullet\nabla_\eta)\,\mathcal{L}\,] = 0$. The thing in brackets is the convective (variational, functional) derivative and its extra pieces arises from that "something new" dependence on ∇η in addition to η . A short-hand for the above is then $\partial_t(\delta_{\dot{\eta}} L) - \delta_\eta L = 0$ and our results look just like our old-time Lagrange Equations, but we have these convective δ/δ things and we use L. You can write this equation of motion in covariant form as $(\nabla_\mu\nabla_\eta{}^\mu)\,\mathcal{L} - \partial_\eta\mathcal{L} = 0$ so we expect our "generalized coordinate" $\eta(x^\mu)$ to be a world scalar, and $\mathcal{L}$ as well. Example of a scalar field.

11.3 Sound vibrations in a gas (355). Here, **η** is the temporary "displacement" of a differential volume of gas from its EQ position. G uses the $\mathcal{L}$ formalism and the adiabatic gas law assumption to derive an equation of motion for **η**. When he uses continuity div **η** = -σ and then uses σ (percentage change in local mass density) as the variable in place of **η**, he gets the 3D wave equation with velocity as in 11-44. This is a good example, we get to turn all the new gears of the formalism including calculation of those convective derivatives.

11.4 The Hamiltonian Formulation (359) We continue with the discrete elastic rod and arrive at a Hamiltonian formalism which is "exactly what you would think". We get Hamilton's equations of motion for example. We have everything in terms of densities: L, T, V, H, π become $\mathcal{L}, \mathcal{T}, \mathcal{V}, \mathcal{H}, \pi$ where the last is the canonical momentum *density* defined in the obvious way. We still have that "something new" element in that we have $\mathcal{H}$(η, π, and also ∇η) and we use the exact same functional derivative method to

handle things. G then redoes the gas problem in the $\mathcal{H}$ formalism and gets the same equation of motion. He then rederives the fact that $dG/dt = [G,H] + \partial G/\partial t$ for G defined as an integral of some density $\mathcal{G}$ and the brackets are the same Poisson Brackets earlier in the book. So if G "commutes" in this sense, it is a constant of the motion.

11.5 Descriptions of Fields (364). Idea that some fields don't need a medium "to undulate in". Full example shows an $\mathcal{L}$ that gives Maxwell's equations. You add matter to this as KE and also as source of ρ and **J** and you get the interaction of matter and light, the basis of QED. A dramatic ending!