Notes for Lecture 18

Coupled Oscillators and Rigid Body

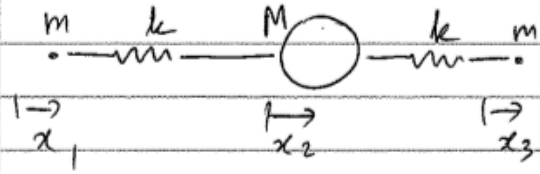
Let us look at couple more examples about coupled oscillators. Then, we cover very basic stuff on rigid body.

18.1 Coupled oscillator example 2

This example is a model for the CO_2 molecule. We are interested in the longitudinal mode only here.

This example is the topic of Ex 12.5 of the textbook. However, notice that the mathematics is somewhat different. Here, we solve a 3×3 problem. We anticipate that we will get one mode for the translation and two modes for vibrations (with the center of mass fixed). This is because the Lagrangian separates into two parts, the motion of the center of mass and the internal motion as viewed from the center of mass coordinate system. The motion of the center of mass is trivial, giving a zero frequency mode. One could have anticipated this and could have reduced the number of generalized coordinates to two instead of three from the beginning (see the textbook). It is your call which method suits you better.

§. Longitudinal modes of a linear molecule (e.g. CO_2)
--- Classical Mechanics model (crude but very useful.)



$$T = \frac{1}{2} m \dot{x}_1^2 + \frac{1}{2} m \dot{x}_3^2 + \frac{1}{2} M \dot{x}_2^2$$

$$U = \frac{1}{2} k (x_1 - x_2)^2 + \frac{1}{2} k (x_3 - x_2)^2$$

$$\overleftrightarrow{M} = \begin{bmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{bmatrix} \qquad \overleftrightarrow{A} = \begin{bmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{bmatrix}$$

$$\left| \overleftrightarrow{A} - \omega_i^2 \overleftrightarrow{M} \right| = 0$$

$$\Rightarrow \left| \frac{\overleftrightarrow{A}}{k} - \frac{\omega_i^2}{k} \overleftrightarrow{M} \right| = 0$$

Let's define $\lambda = \frac{\omega_i^2}{k} m$

$$\begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\frac{M}{m}\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)\left\{\left(2-\frac{M}{m}\lambda\right)(1-\lambda)-1\right\} + \left\{(-1)\cdot(1-\lambda)\right\} = 0$$

$$(1-\lambda)\left\{\left(2-\frac{M}{m}\lambda\right)(1-\lambda)-2\right\} = 0$$

$$(1-\lambda)\cdot\left\{\frac{M}{m}\lambda^2 - \left(2+\frac{M}{m}\right)\lambda\right\} = 0$$

$$\lambda = 1, \; 0, \; \frac{M+2m}{M} \quad \checkmark$$

① $\lambda = 0 \dashrightarrow \omega_1^2 = 0$ $\quad x_1 = x_2 = x_3$ $\quad \vec{T}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0 \quad \rightarrow \quad \begin{array}{l} u = v \\ v = \frac{1}{2}(u+w) \\ v = w \end{array} \quad \rightarrow \quad u = v = w$$

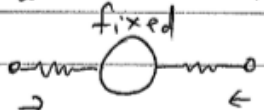
$\vec{T}_1$

translation!!

no restoring force!

② $\lambda = 1 \dashrightarrow \omega_2^2 = \frac{k}{m}$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 2-\frac{M}{m} & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0 \quad \rightarrow \quad \begin{array}{l} v = 0 \\ u = -w \end{array} \qquad \vec{T}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$



$\vec{x}_1$ $\qquad -\vec{x}_1$ "breathing" mode

$$\vec{x} = \vec{T}_2 \eta_2 = \begin{bmatrix} \eta_2 \\ 0 \\ -\eta_2 \end{bmatrix}$$

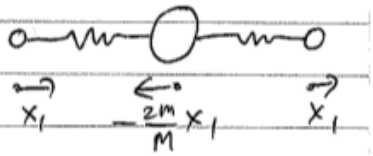
$\therefore x_1 = -x_3, \quad x_2 = 0$

③ $\lambda = \frac{M+2m}{M} \dashrightarrow \omega_3^2 = \frac{M+2m}{Mm} \cdot k$

$$\begin{bmatrix} -\frac{2m}{M} & -1 & 0 \\ -1 & -\frac{M}{m} & -1 \\ 0 & -1 & -\frac{2m}{M} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0 \quad \rightarrow \quad \begin{array}{l} v = -\frac{2m}{M} u \\ v = -\frac{2m}{M} w \end{array} \quad \Rightarrow \quad u = w$$

$$\vec{T}_3 = \begin{bmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{bmatrix}$$

$$\vec{x} = \vec{T}_3 \eta_3 = \begin{bmatrix} \eta_3 \\ -\frac{2m}{M}\eta_3 \\ \eta_3 \end{bmatrix}$$



$x_1 = x_3 \qquad x_2 = -\frac{2m}{M} x_1$

meaning?

x_1 and x_3 in phase

while CM (center of mass) $= \frac{m(x_1+x_3)+x_2 M}{2m+M} = \frac{2mx_1 - 2mx_1}{2m+M} = 0$

Now, to complete the solution, we ask

what is $\vec{\eta} = ?$

We need to find $\overleftrightarrow{T}^{-1}$. A little tricky...

$$\overleftrightarrow{T} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -\frac{2m}{M} \\ 1 & -1 & 1 \end{bmatrix} \Leftarrow \text{Cannot be made into an orthogonal matrix!}$$

$\vec{T}_1 \quad \vec{T}_2 \quad \vec{T}_3$

$\vec{T}_3 \cdot \vec{T}_1 \neq 0$

We have to invert $\overleftrightarrow{T}$ by brute force...

$$\overleftrightarrow{T}^{-1} = (-)\cdot\frac{M}{2M+4m}\cdot\begin{bmatrix} -\frac{2m}{M} & -2 & -\frac{2m}{M} \\ -1-\frac{2m}{M} & 0 & \frac{2m}{M}+1 \\ -1 & 2 & -1 \end{bmatrix}$$

$$\Rightarrow \begin{cases} \eta_1 = \frac{1}{M+2m}\left(mx_1 + Mx_2 + mx_3\right) \\ \eta_2 = \frac{1}{2}\left(x_1 - x_3\right) \\ \eta_3 = \frac{M}{2(M+2m)}\left(x_1 + x_3 - 2x_2\right) \end{cases}$$

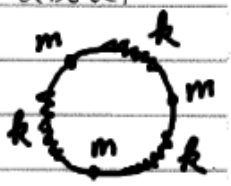
↖ CM coordinate

Note

① Why was it not possible to turn $\overleftrightarrow{T}$ into an orthognal matrix by normalizing each column vector?

Ans. $\overleftrightarrow{M}$ was not a constant times an identity matrix.

② In some problems, λ (or ω_i^2) may be degenerate --- what to do about $\vec{T}_C$ in those cases?



Ans. Indeed, this happens. E.g.,

$$\overleftrightarrow{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \overleftrightarrow{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

up to a const up to a const

$\Rightarrow \lambda = 1$ or 4

$\rightarrow$ degenerate (double root)

$\Rightarrow$ For $\lambda = 1$, all we get for $\vec{T}_C = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$ is

$$u + v + w = 0.$$

It is a plane!

You can pick two orthogonal vectors in that plane, e.g., $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ as $\vec{T}_1$ and $\vec{T}_2$

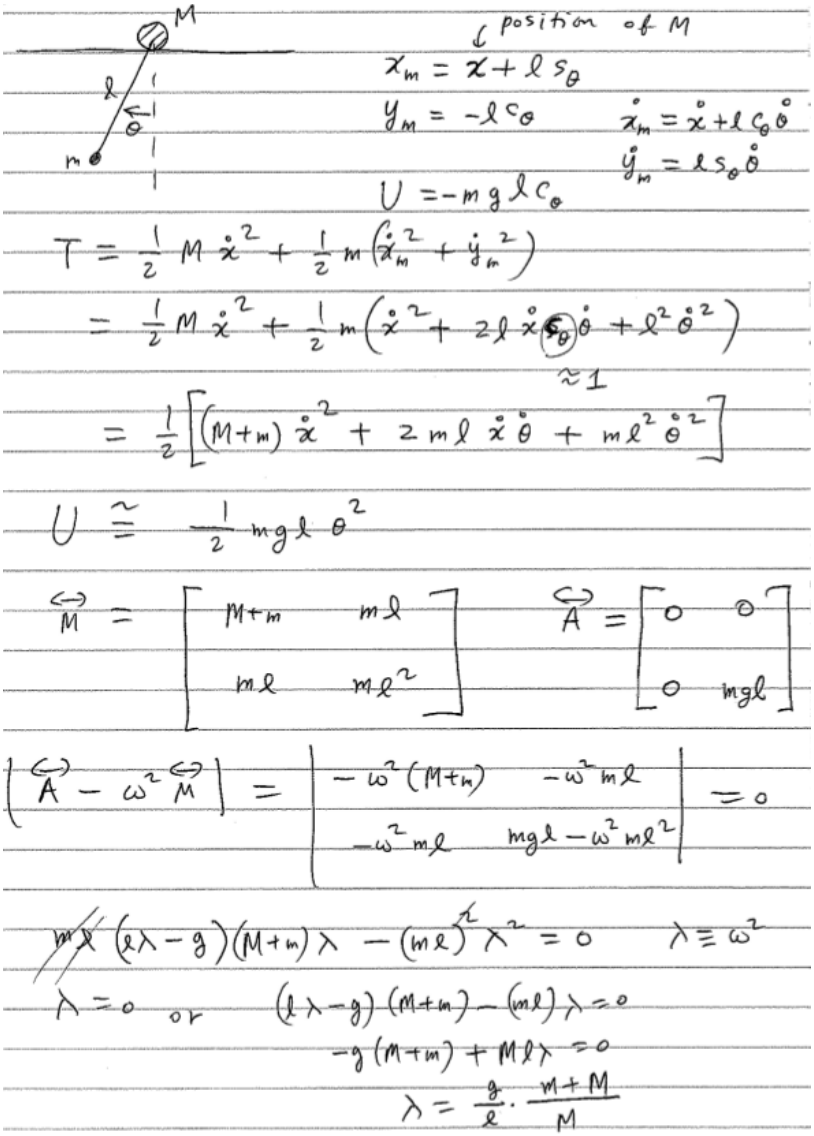
You are free to choose any two linearly independent vectors for a doubly degenerate eigenvalue.

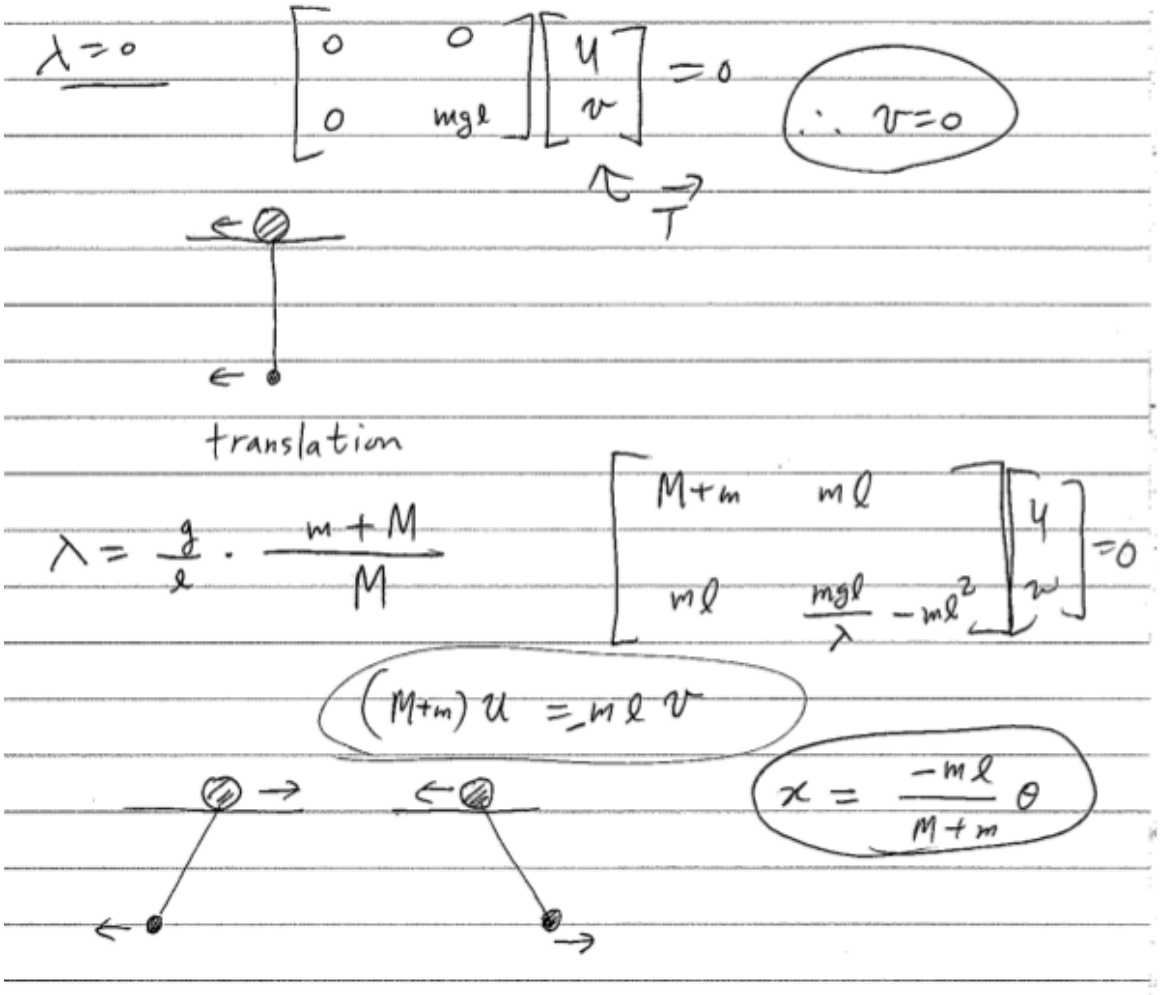
18.2 Coupled oscillator example 3

This example shows that the coupling does not necessarily occur through the potential energy term. It can occur through the kinetic energy term as well.

In this example, a mass is moving on a smooth rail, while a simple pendulum is connected to it.

Observe again that the two modes are: one translation mode and the other an oscillation mode with the center of mass fixed.



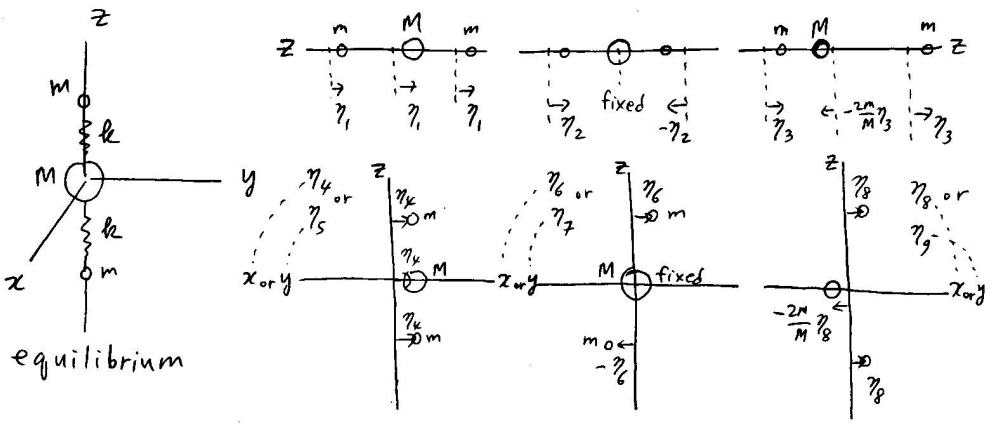


18.3 Normal modes of a molecule

Above, we studied a classical mechanics model[1] for CO_2. By choice, we only considered one dimensional motions in which any atom moves only along the line of the molecule. We obtained one translational mode (zero eigen-frequency) and two vibrational modes (finite eigen-frequencies). Of course, we should expect other modes, as each atom can move in all directions, not just along the line of the molecule. Here is

[1] The correct quantum mechanical treatment of these molecules can be made by "simply" (what I mean is that it is simple once your learn quantum mechanics) quantizing each normal coordinate η_i after, as we discussed in the last lecture, conveniently splitting the Lagrangian, and thus the Hamiltonian, into a sum of contributions from normal modes. So, the exercise of classical coupled oscillator model for molecules is a very rewarding one, as it lays a solid foundation for the correct quantum mechanical treatment.

a sketch of all modes.



In the above diagram, some modes $(\eta_1, \ldots, \eta_3)$ are those we already considered. As with the η_1 mode, some modes turn out to be zero frequency oscillation – which means that it is not really an oscillation – but that is fine. Let us analyze the whole situation here, by asking and answering some questions.

1. Why do we have nine modes in total? Ans: because the total degrees of freedom $= 3 \times 3 = 9$.

2. What modes correspond to no restoring force, i.e. zero eigen-frequency? Ans: $\eta_1, \eta_4, \eta_5, \eta_6, \eta_7$.

3. What is the nature of each zero frequency mode? Ans: translation (η_1, η_4, η_5) and rotation (η_6, η_7).

4. Which modes are degenerate (i.e. of the same eigen-frequency)? Ans: All zero frequency modes are degenerate. The pair η_6, η_7 are degenerate, as are the pair η_8, η_9.

5. Which modes have a finite total momentum? Ans: all translational modes (η_1, η_4, η_5) only.

6. Which modes have a finite total angular momentum? Ans: All rotational modes (η_6, η_7) and all translational modes (η_1, η_4, η_5). The latter, only if the angular momentum is measured around a point that is chosen to be not a special point (i.e. not the center of mass).

7. Why is it that η_6, η_7 modes have no restoring force but η_8, η_9 modes have restoring force (and thus a finite frequency)? Ans: This is an excellent question.

Notice that in terms of the spring force argument all these modes $(\eta_6, \ldots, \eta_9)$ should have zero restoring force, since the length change of the spring is 2nd order, not 1st order, in η. So, there is more physics than just Hooke's laws involved with the two springs here. Microscopically, the origin can be understood as a generation of the transverse dipole moment for the η_8, η_9 modes, while no such dipole moment develops for the η_6, η_7 modes. This is the reason why there is a finite restoring force for the η_8, η_9 modes, but zero restoring force for the η_6, η_7[2]. For this reason, the potential energy for the transverse displacements x_1, x_2 and x_3 must be written as $U \propto [(x_1 - x_2) + (x_3 - x_2)]^2$.

Having done this exercise, it may not be too difficult to understand the following table, which is valid for a molecule consisting of any number of atoms (tot-DOF = total degrees of freedom).

shape of molecule	Number of masses (atoms)	tot-DOF	translational modes	rotational modes	vibrational modes
line	n	$3n$	3	2	$3n-3-2$
non-line	n	$3n$	3	3	$3n-3-3$

The only difference for the two types (line or non-line) of molecules is the number of vibrational modes. It is meaningless to talk of a rotation of a linear molecule around its own axis – this is why a linear molecule has only two distinct rotations (η_6, η_7 of the above example), while a non-linear molecule has rotation modes in all three directions.

Also, note that for a linear molecule (e.g.CO_2, H_2), one can distinguish between longitudinal modes (for which atoms move along the line of the molecule) and transverse modes (for which atoms move perpendicular to the line of the molecule). For a linear molecule with n atoms, there are $n-1$ longitudinal modes (the total of n degrees of freedom within that one dimensional space minus 1 for translation). Thus, the total number of transverse vibrational modes is $3n-3-2-(n-1) = 2n-4$. It is no coincidence that this is an even number. It is for the following reason. Suppose there is a transverse mode where the displacements of all atoms are in the x direction, while the linear molecule is aligned along the z axis. Then this transverse mode has a twin mode where the displacements of all atoms, while identical in value, are in the y direction. These two modes must both exist, and they have the same eigen-frequency, i.e. they are degenerate. For either direction (x or y) of displacement, then, there will

[2]However, all this discussion is in the spirit of the *Harmonic approximation*. See later discussions in this note.

be $n-2$ normal modes. Indeed, in the above example of CO_2, there was only one transverse mode per direction.

Knowing normal modes before solving matrices

If all normal modes are obvious from the get-go, then in principle the following is all one needs to find all normal modes. Write down all normal modes (namely all $\vec{T}_i$ vectors, as pictorially given in the above diagram), and then prove, for each normal mode, that *each* non-stationary mass vibrates at the same frequency. For the six diagrams shown above, it is rather easy to do this! [Please try this!] For instance, for the 3rd digram, one can show easily that *each* of all three masses should vibrate at the angular frequency $\frac{(1+2m/M)k}{m}$, proving that this is indeed a normal mode with that angular frequency.
So, here is a question of great physical and practical importance. How can one know as much as possible about all these normal modes even before beginning to write down $\overleftrightarrow{A}$ and $\overleftrightarrow{M}$? A few guides are useful.

1. Distinguish between modes that have a finite total momentum (i.e. the center of mass is moving) and those that don't (i.e. the center of mass is fixed). The modes that do are translational modes.

2. Then, among the modes that have zero momentum, distinguish between modes that have a finite angular momentum and those that don't. Those that do are rotational modes.

3. Use "symmetry." If two masses are identical, then they can be swapped without changing physics. Then, the displacements of those two masses (say, x_1 and x_2) will appear as either the same ($x_1 = x_2$) or the opposite ($x_1 = -x_2$) in normal modes. Also, the "symmetry" is the reason why the longitudinal modes and the transverse modes can be considered separately from the get-go.

18.4 Normal modes, general remarks

In the last two lectures and in the previous section, we considered a coupled oscillator problem assuming small oscillations. This elementary treatment generally goes

by the name of "**harmonic approximation**." The "stars" of this treatment were definitely those things called **normal modes**. In the harmonic approximation of coupled oscillators, all potential energies and all kinetic energies are taken only up to the quadratic terms of generalized coordinates, and the result is that the system can be described in terms of completely independent normal modes, each of which does a simple harmonic oscillation.

It should seem both satisfying and unrealistic to have such a nice picture. It is very satisfying in the sense that there is a certain beauty to this separation of the whole complex problem into completely independent and easy bits. At the same time, you might mutter to yourself that this is a bit artificial in that any small perturbation (cubic or higher terms ignored in the Lagrangian that will introduce the "non-linearity") would spoil this ideal picture of completely independent normal modes. Indeed. Like any physics models, the harmonic approximation of coupled oscillators is a bit simplistic but immensely useful *starting point* for describing real world phenomena. For any given stable object, a small deformation will give rise to a restoring force that, in the first approximation, follows Hooke's law, and so the harmonic approximation is of relevance to *any* object.

We saw that some normal modes were "trivial." Like that of translation in which all masses move by the same amount. Or, rotation. We can call these **shape-and-size-preserving normal modes**. They are translations and rotations[3]. You take a snap shot of the system for which only translational or rotational normal modes are excited, and you will find that the system looks just the same in every detail at any time, except for the overall position and the overall orientation. These translational and rotational modes involve no restoring force, and thus are characterized by zero eigen-frequencies. Then, there are **shape-or-size-changing normal modes**. These are all the rest – vibrations, we call them. Of course, vibrations do not change the average shape of the system, within the harmonic approximation. At any given time, though, a snapshot of the system will show a distorted shape of the system.

Let us consider the following two completely different questions. First, would there be such a case when vibrational normal modes can be completely ignored? Second, would there be such a case when vibrational normal modes must be treated definitely beyond the harmonic approximation? The answers are yes and yes.

The first case is the case of a "**rigid body**." In this case, the vibrational modes

[3]Perhaps there is a subtle point to keep in mind here. Strictly speaking all modes that we have been considering are small "oscillations." Of course, if the oscillation frequency is zero, then it is no longer an oscillation! The mode corresponds to a small part of translation or rotation. In the case of translation, one can simply repeat small translations to get the finite translation. For the rotation though, this is not possible. No matter, the normal modes corresponding to translation and rotation represent infinitesimal motions that are equivalent translation and rotation.

are deemed unimportant and are ignored, in comparison to the translational and rotational modes. This is an important topic that we will discuss below.

The second case, while being more difficult to discuss, is also quite interesting and important nonetheless. To appreciate this second case, consider the following statement. **Within the harmonic approximation, a coupled oscillator system will always remain "brand-new"**. As any object can be thought of as a coupled oscillator system to the first approximation, then this statement means the following. **No object will ever wear out, if it can be completely modeled as a harmonic coupled oscillator system**. Why so? In this nice picture of independent simple harmonic normal modes, all that can happen to the object in a collision, or any physical process, is the increase or decrease of energy of each normal mode. So, the object can, for example, vibrate more or less but on the average its shape or size can never change. In other words, it will never lose its average form – it never wears out! Our experience is quite the opposite – things wear out especially if they are subject to violent collision processes (baseball hit by a baseball bat, tennis ball hit by a racket, a tire rubbed against the road surface, etc). So, it must be the case that non-linear terms play an essential role in normal *wear and tear* processes. Indeed, without non-linear terms in the Lagrangian, it would be impossible to explain how things can (slightly) deform after a collision or why some materials can be lost during a collision.

18.5 Rigid body – definition

A rigid body was defined already in Section 18.4. Here it is. **A rigid body is a system of particles where all vibrational modes can be ignored compared to translational and rotational modes.** This means that there are only six degrees of freedom to consider for a rigid body, greatly simplifying the mathematics required.

The standard definition of a rigid body can be stated as followed. **A rigid body is a collection of particles, for which the distance of between any pair of particles is fixed.**

This definition simply means that a rigid body is an object whose shape and size are considered fixed. Then, the only possible excitation of such a body is the shape-and-size-conserving normal modes, i.e. translations and rotations. This makes the standard definition equivalent to the first definition.

[Subtle advanced discussion here; can skip if you like] Strictly speaking, in truth,

all normal modes of every object *always* have finite amplitudes.[4] However, if one is considering cases in which an object is given a total momentum or a total angular momentum that is very large so that the translational or the rotational modes of that object have much greater amplitude than all vibrational modes, then it is of little importance to consider to those vibrational modes as far as the macroscopic dynamics of the body is concerned. Thus, a notion of a rigid body.

When to apply the rigid body concept and when not. One must keep in mind that, the concept of rigid body fails in general during the short collision process. Before or after a collision, the concept of rigid body is useful. One may be well-advised, though, that even when the body "looks the same" before and after the collision, the internal state may have changed during the collision process. If it did, then we have an inelastic collision. If it did not, then, assuming that the energy was not lost to heat or sound during the collision, we have an elastic collision. In a similar vein, it is important to note that the concept of rigid body is useful even when the body is really not rigid in general. For instance, consider a figure skater spinning on ice, with arms tugged in initially and, then, arms stretched. In a typical problem to figure out how the angular velocity changes, we learn to deal with the body of the figure skater as though it was a rigid body in the initial state and another rigid body in the final state. This, despite the fact that a human body is generally not "rigid" as it is flexible to change. It is nevertheless correct and useful to consider a human body a rigid body in a small time window when the shape and size of the body does not change. Of course, during the stretching of her arms, the concept of rigid body is quite unsuitable to the whole body.[5]

18.6 Rigid body – general properties

In Section 14.2, we have laid out some important properties of $\vec{P}, \vec{L}, T$ – the total momentum, the total angular momentum, and the total kinetic energy – of a general system of particles.

Here we shall use the same notation in that section.

[4] This is because at a finite temperature the thermal energy excite them. At $T = 0$, however, all amplitudes will cease to exist, if classical mechanics were to hold at such a low temperature. However, at $T = 0$, classical mechanics is not valid, and all normal modes have finite amplitudes. This is due to the Heisenberg uncertainty principle. We also use the expression "quantum fluctuations" to describe this $T = 0$ physics. In other words, things are agitated due to "thermal fluctuations" at a finite T and "quantum fluctuations" at zero T.

[5] Even then, (approximately) shape-and-size conserving *parts* of bodies can be considered rigid bodies.

Let us discuss these quantities for a rigid body. To specify the motion of a rigid body, one only needs to note the following.

Translation	Described by the change of $\vec{R}$. This carries the total momentum, $\vec{P} = M\dot{\vec{R}}$.
Rotation	Each coordinate $\vec{r}'_\alpha \equiv \vec{r}_\alpha - \vec{R}$ rotates around $\vec{R}$ at the same rate at any given time.

For a rotation by $d\vec{\phi}$, we have

$$d\vec{r}'_\alpha \;=\; d\vec{\phi} \times \vec{r}'_\alpha$$

For the derivation of this relation, see LN 9, page 6. Note that there we used the δ notation, $\delta\vec{\phi}$ and $\delta\vec{r}$, as we were considering a virtual rotation (fixed time). Here, we are using the notation $d\vec{\phi}$, suitable for a real dynamical rotation of a rigid body. In any case, dividing the above equation by dt, we get

$$\dot{\vec{r}}'_\alpha \;=\; \vec{\omega} \times \vec{r}'_\alpha$$

where $\vec{\omega} \equiv d\vec{\phi}/dt$ is the angular velocity around the center of mass. This is a key relation for describing the motion of a rigid body. It is also important for any problem involving a rotating frame.

Let us examine the kinetic energy of a rigid body. Let us examine the internal part, T', the kinetic energy of particles as measured in the center of mass frame.

$$\begin{aligned}
T' &= \frac{1}{2}\sum_\alpha m_\alpha \dot{\vec{r}}'_\alpha \cdot \dot{\vec{r}}'_\alpha \\
&= \frac{1}{2}\sum_\alpha m_\alpha (\vec{\omega} \times \vec{r}'_\alpha) \cdot (\vec{\omega} \times \vec{r}'_\alpha) \\
&= \frac{1}{2}\sum_\alpha m_\alpha \left(\omega^2 {r'_\alpha}^2 - \sum_{j,k=1}^{3} \omega_j r'_{\alpha,j} \omega_k r'_{\alpha,k} \right) \\
&= \frac{1}{2}\sum_{j,k=1}^{3} \sum_\alpha m_\alpha \left(\omega_j \omega_k \delta_{jk} {r'_\alpha}^2 - \omega_j r'_{\alpha,j} \omega_k r'_{\alpha,k} \right) \\
&= \frac{1}{2}\sum_{j,k=1}^{3} \omega_j I_{jk} \omega_k
\end{aligned}$$

where the **inertia tensor** I_{jk} is defined as

$$I_{jk} \;=\; \sum_\alpha m_\alpha \left(\delta_{jk} {r'_\alpha}^2 - r'_{\alpha,j} r'_{\alpha,k}\right)$$

Here, again, in the expression of the kinetic energy, we see a quadratic form, as in the normal mode problem,

$$T' \;=\; \frac{1}{2} \vec{\omega}^t \overleftrightarrow{I}\, \vec{\omega}$$

Now, let us consider the angular momentum vector. Again, our concern is $\vec{L}'$ only, corresponding to the internal motion viewed in the center of mass frame, as $\vec{L}_M$ is simple already.

$$\begin{aligned}
\vec{L}' \;&=\; \sum_\alpha m_\alpha \vec{r}'_\alpha \times \dot{\vec{r}}'_\alpha \\
&=\; \sum_\alpha m_\alpha \vec{r}'_\alpha \times (\vec{\omega} \times \vec{r}'_\alpha) \\
&=\; \sum_\alpha m_\alpha \left(\vec{\omega} {r^2}'_\alpha - \vec{r}'_\alpha (\vec{\omega} \cdot \vec{r}'_\alpha)\right) \\
&=\; \overleftrightarrow{I}\, \vec{\omega}
\end{aligned}$$

with the same inertia tensor as defined above[6]. To summarize,

$$\vec{L}' \;=\; \overleftrightarrow{I}\, \vec{\omega}$$

18.7 Rigid body – inertia tensor

It helps to rewrite the inertia tensor in a more familiar form.

[6]In this derivation, the main vector identity used is $\vec{A} \times (\vec{B} \times \vec{C}) \;=\; \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$, often referred to as a "bac cab" identity. The more fundamental identity is the following famous identity, which is the underlying explanation for this derivation and the previous derivation of T': $\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{ilm} \;=\; \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$, where ε is the Levi-Civita symbol and δ is the Kronecker delta symbol.

Recall that the vector $\vec{r}'_\alpha$ is the position vector of mass m_α in the CM frame. Let us write its components $r'_{i,j}$ as $r'_{i,1} = x'_\alpha$, $r'_{\alpha,2} = y'_\alpha$, and $r'_{\alpha,3} = z'_\alpha$, keeping in mind that these are measured relative to the center of mass, $\vec{R}$. The inertia tensor is then

$$\overleftrightarrow{I} = \sum_\alpha \begin{pmatrix} m_\alpha({y'_\alpha}^2 + {z'_\alpha}^2) & -m_\alpha x'_\alpha y'_\alpha & -m_\alpha x'_\alpha z'_\alpha \\ -m_\alpha x'_\alpha y'_\alpha & m_\alpha({x'_\alpha}^2 + {z'_\alpha}^2) & -m_\alpha y'_\alpha z'_\alpha \\ -m_\alpha x'_\alpha z'_\alpha & -m'_\alpha y'_\alpha z'_\alpha & m_\alpha({x'_\alpha}^2 + {y'_\alpha}^2) \end{pmatrix}$$

The matrix given here is for a point mass m_α and the summation is made over matrices. The summation of matrices mean that the inertia tensor is additive for different parts of the body. Of course, the summation symbol can be brought into the matrix – each component of the matrix should then be summed over. If the mass distribution is continuous, then $\sum_\alpha m_\alpha \to \int dm$.

$$I_{jk} = \int dm \left(\delta_{jk} r'^2 - r'_j r'_k\right)$$

One can recognize that the diagonal terms are what one would expect from the definition given in introductory classes of mechanics: the rotational inertia for a point mass is the mass times distance squared, with distance here meaning the distance between the mass and the rotational axis. For instance I_{11} corresponds to the rotational inertia around the x' axis, and so ${y'_\alpha}^2 + {z'_\alpha}^2$ in the expression for I_{11} should be recognized as the square of the distance to the x' axis.

Notice that $\overleftrightarrow{I}$ is a real symmetric matrix. This means that, by the well-known theorem in Linear Algebra, it can be diagonalized by an orthogonal matrix $\overleftrightarrow{O}$:

$$\overleftrightarrow{O}^t \, \overleftrightarrow{I} \, \overleftrightarrow{O} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}$$

The three column vectors $\vec{O}_j$ of the matrix $\overleftrightarrow{O}$ are, then, the eigenvectors of the inertia tensor $\overleftrightarrow{I}$ with eigenvalue I_j:

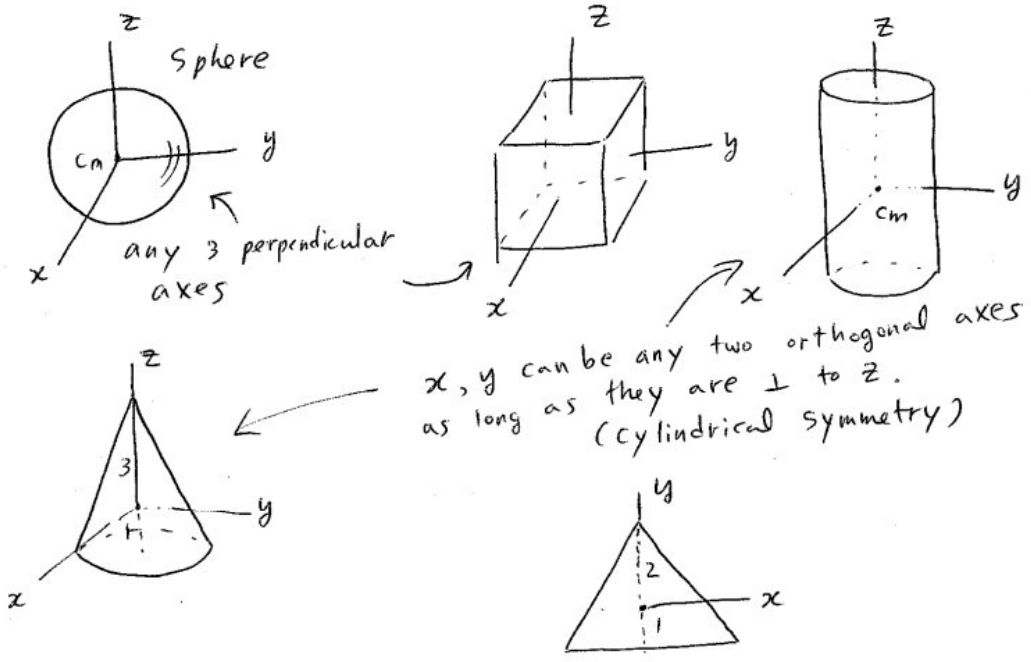
$$\overleftrightarrow{I} \, \vec{O}_j = I_j \vec{O}_j$$

The three orthogonal axes corresponding to the three vectors $\vec{O}_j$ are called the **principal axes of inertia** for the rigid body. The three eigenvalues I_j's are called the **principal moments of inertia**, as opposed to just the **moment of inertia**, which refers to a diagonal element, I_{jj}, of any inertia tensor, a diagonal matrix or not.

Note that the above property of diagonalizability applies to an object with *any* shape. An object may be shaped like a leaf of a tree, a pumpkin or a pebble. However irregular or regular an object's shape is, it is always possible to find three mutually orthogonal axes so that the inertia tensor is diagonal.

For an object with a certain symmetry, it is usually intuitively obvious how to take the principal axes of inertia. The figure below shows some typical examples. If the object is very symmetric (like a sphere, a cube, a cylinder, etc.), then there can be many, often infinite, ways, to take the principal axes of inertia.

[Now, the prime in x' etc. is dropped for simplicity, for the rest of this section.]



For a planar system of particles, it is obvious that a principal axis (call it z) must be perpendicular to the plane, and the other two principal axes (call them x, y) in the plane. Then, for all masses, $z_i = 0$. Therefore, $I_1 = \sum_i m_i(y_i^2 + z_i^2) = \sum_i m_i y_i^2$, and $I_2 = \sum_i m_i(x_i^2 + z_i^2) = \sum_i m_i x_i^2$, while $I_3 = \sum_i m_i(x_i^2 + y_i^2)$.

For a planar system of particles distributed in the xy plane, the following property holds, assuming that xyz axes are the principal axes of inertia.

$$I_3 \;=\; I_1 + I_2$$

When the two principal moments of inertia are identical, then we talk of a "**symmetric top**" or a "**symmetric rigid body**." When the three principal momenta of inertia are identical, then we have a case of "**spherical top**." This is not only the case of a sphere, but also the case of a cube.

18.8 Rigid body – pure rotation around any point

So far, we worked strictly within the center of mass frame only, when we discussed a rigid body.

The center of mass frame is of fundamental importance. As we have seen in Section 14.2, the total linear momentum, the total angular momentum, or the total kinetic energy of a system can *always* be split nicely into two terms, the contribution of a point particle of the total mass M at the position vector $\vec{R}$ and the contribution as viewed from the center of mass reference frame. [For the linear momentum, the latter contribution is always zero, by definition.] This nice separation would not be possible, in general, if any point other than the center of mass is chosen.

However, in some cases, it is useful to consider rotation of a rigid body around a reference point that is *not* the center of mass. This would be the case if a well-chosen non-center-of-mass reference point makes the description of the whole motion simpler rather than more complicated. When can such a simplification occur? An example is when the reference point can be chosen so that the entire motion can be described as a pure rotation. For instance, the motion of a rolling wheel is a pure rotation around the instantaneous contact point. A physical pendulum is another example. It is much simpler to describe the physical pendulum problem as a pure rotation around the pivot point rather than as a motion of the center of mass plus a rotation of the object around its center of mass.

So, let us consider **a pure rotation of a rigid body around a fixed arbitrary point** Q.

For such a pure rotation, the following relations are derived in exactly the same way as we did for the center of mass frame.

$$\vec{L} = \overleftrightarrow{I}_Q \, \vec{\omega}_Q$$
$$T = \frac{1}{2} \, \vec{\omega}_Q^{\,t} \, \overleftrightarrow{I}_Q \, \vec{\omega}_Q$$

Here, we use the subscript Q to note that the inertia tensor and the angular frequency are relative to the point Q. Note that, since we assume a pure rotation, $\vec{L}$ and T are now the *total* angular momentum and the *total* kinetic energy, respectively.

Here, the inertia tensor $\overleftrightarrow{I}_Q$ is given by the exactly same formula as in the previous class, but the position vector $\vec{r}'_\alpha$ (which was in the center of mass reference frame) is now replaced by $\vec{r}_{\alpha Q}$ (which is now referenced to point Q):

$$I_{Q,jk} \;=\; \sum_\alpha m_\alpha \left(\delta_{jk} r_{\alpha Q}^2 - r_{\alpha Q,j} r_{\alpha Q,k}\right)$$

In the above expression for the inertia tensor, we can use $r_{\alpha Q,j} = R_{Q,j} + r'_{\alpha,j}$, where $\vec{R}_Q$ is the position vector of the center of mass relative to Q, to find the relationship between the inertia tensors $\overleftrightarrow{I}_Q$ and $\overleftrightarrow{I}$, the latter the inertia tensor in the center of mass reference frame as we discussed in the previous section.

$$\begin{aligned}
I_{Q,jk} \;&=\; \sum_\alpha m_\alpha \left(\delta_{jk} r_{\alpha Q}^2 - r_{\alpha Q,j} r_{\alpha Q,k}\right) \\
&=\; \sum_\alpha m_\alpha \left(\delta_{jk}\left[\sum_l (R_{Q,l} + r'_{\alpha,l})^2\right] - (R_{Q,j} + r'_{\alpha,j})(R_{Q,k} + r'_{\alpha,k})\right) \\
&=\; \sum_\alpha m_\alpha \left(\delta_{jk}(R_Q^2 + r_\alpha'^{\,2}) - R_{Q,j} R_{Q,k} - r'_{\alpha,j} r'_{\alpha,k}\right) \qquad \because \sum_\alpha m_\alpha r'_\alpha = 0 \\
&=\; I_{jk} + M(\delta_{jk} R_Q^2 - R_{Q,j} R_{Q_k})
\end{aligned}$$

This is a very important result, going by the name of the **parallel axis theorem**. The parallel axis theorem is as yet another example of the "nice separation" for a physical quantity, when the center of mass frame is utilized. Namely, the total inertia tensor is nicely separated into the sum of two terms: (1) the inertia tensor for a point mass M at the position of the center of mass and (2) the "internal" inertia tensor of the body with respect to the center of mass. Such a nice separation occurred for the linear momentum, the angular momentum, or the kinetic energy, as well.

Parallel axis theorem

The inertia tensor $\overleftrightarrow{I}_Q$, defined with respect to an arbitrary point Q, is related to the inertia tensor $\overleftrightarrow{I}$, defined with respect to the center of mass, by the following relation, where $\vec{R}_Q$ is the position vector of the center of mass with respect to point Q.

$$I_{Q,jk} \;=\; M(\delta_{jk}R_Q^2 - R_{Q,j}R_{Q,k}) + I_{jk}$$

or

$$\overleftrightarrow{I}_Q \;=\; M\left(R_Q^2 - \vec{R}_Q\vec{R}_Q^t\right) + \overleftrightarrow{I}$$

In the second form, the term R_Q^2 means a unit matrix times that term. Also, note that $\vec{R}_Q$ is a column vector in the second form, and thus $\vec{R}_Q\vec{R}_Q^t$ is a square matrix (it is the direct product of $\vec{R}_Q$ with itself).