Notes for Lecture 5

Driven SHO and non-linear oscillations

5.1 Driven SHO

(This part is almost repeated from the old version of lecture note 4, but it has small modifications here and there and a very large change in section 5.1.3. This version recommended very highly!)

5.1.1 Particular solution, complementary function

Suppose that a force, $F(t)$, drives a SHO.

$$\begin{aligned} m\ddot{x} + b\dot{x} + kx &= F(t) \\ \ddot{x} + 2\beta\dot{x} + \omega_0^2 x &= f(t) && f(t) \stackrel{def}{=} F(t)/m \\ Lx &= f(t) && L \stackrel{def}{=} \frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2 \text{ is a } \textit{linear operator} \text{ acting on } x(t). \end{aligned}$$

Now, we have a so-called **"inhomogeneous" differential equation**, or a **differential equation with a "source."** The $f(t)$ function is the source term.

No matter, we should remember that the general solution to this ODE is the one that has two integration constants (2 times 1 as the degree of freedom is 1). If we manage to find such a solution, then by the uniqueness of the solution of Newton's equation, that is *the* general solution.

How do we find them? The answer is that we already know a lot about the above equation in its homogeneous form ($f(t) = 0$). The general solution for the homogeneous equation is called a **complementary function**, $x_c(t)$. Note that the complementary solution already contains two integration constants to adjust to any initial condition specified. What does this mean? **All we have to do is then to find one particular solution to the above equation.** Call that particular solution $x_p(t)$. This particular solution should not, and need not, have any integration constant.[1]

Assume that we have found both $x_c(t)$ and $x_p(t)$. Then, the proof that $x(t) = x_c(t) + x_p(t)$ is the general solution is pretty simple, since L, as defined above, is a linear operator.

$$L(x_c + x_p) \;=\; Lx_c + Lx_p \;=\; 0 + f(t) \;=\; f(t) \qquad\qquad \text{QED.}$$

Let me remind you **what it means that L is a linear operator**. It means that for any numbers a, b and any functions $x_1(t), x_2(t)$,

$$L(ax_1(t) + bx_2(t)) \;=\; aLx_1(t) + bLx_2(t)$$

(see pages 9, 14, and 4 of lecture note 1, also). For the current SHO problem, $L = \frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2$ is definitely a linear operator (cf. homework), due basically to the distributive rule of the differential operator $d(f+g)/dt = df/dt + dg/dt$.

A general approach that will give *any* particular solution x_p will be worked out by you in homework. Here, we use a more elementary approach. For either approach, though, the following principle forms the foundation.

5.1.2 Superposition principle

A **linear system** displays this important principle. It means the following.

Suppose we have a system defined by the following equation

$$Lx(t) \;=\; f(t)$$

where L is a **linear operator** on function $x(t)$, and $f(t)$ is a source term for this equation. Further, suppose that one can divide the source term into two terms, $f(t) = f_1(t) + f_2(t)$, and that

$$Lx_1(t) \;=\; f_1(t) \;\text{ and }\; Lx_2(t) \;=\; f_2(t)$$

[1] Should you be writing the particular solution with some extra integration constants, that means you are repeating some part of the complimentary function in the particular solution. That is of course redundant.

Then, the superposition principle means that $x_1(t) + x_2(t)$ is the solution for the original equation.

$$L(x_1(t) + x_2(t)) \;=\; f_1(t) + f_2(t) \;=\; f(t)$$

This is easy to prove since $L(x_1+x_2) = Lx_1+Lx_2$, due to the linearity of the operator L.

You will note that we made use of this linear property already! Indeed, we already used the superposition principle! We just did not use this fancy name. For example, the solution for the homogeneous ODE was written as a linear combination of two independent solutions. This was the superposition principle, with $f_1 = f_2 = 0$. When we discussed the general solution for the driven SHO problem in the last section, that was the superposition principle, with $f_1 = 0$ (for the complementary solution) and $f_2 = f(t)$ (for the particular solution).

Notice that the superposition principle can be immediately extended to *any number of components* into which $f(t)$ can be decomposed: $f(t) = \sum_i f_i(t)$ or $f(t) = \int d\alpha \tilde{f}(\alpha, t)$. If each component gives rise to a certain solution ($f_i(t) \to x_i(t)$, $\tilde{f}(\alpha, t) \to \tilde{x}(\alpha, t)$), then the solution to the problem is

$$x(t) \;=\; \sum_i x_i(t) \qquad\qquad \text{for } f(t) = \sum_i f_i(t)$$

$$x(t) \;=\; \int d\alpha\; \tilde{x}(\alpha, t) \qquad\qquad \text{for } f(t) = \int d\alpha \tilde{f}(\alpha, t)$$

The functions $f_i(t)$ or $\tilde{f}(\alpha, t)$ can be any functions. Typical examples are when these are Fourier expansions of the original function.[2]

Physically the superposition principle means that each solution $x_i(t)$ or $\tilde{x}(\alpha, t)$ remains intact even when it is combined with all other solutions. Individual solutions "just add up." Note that, in this addition, what is added is the amplitude ($x_i(t)$ or $\tilde{x}(\alpha, t)$), not the intensity ($|x_i(t)|^2$ or $|\tilde{x}(\alpha, t)|^2$). This is the essential feature of the superposition, important for understanding the "interference phenomena" for light and other waves.

Indeed, the superposition property is an essential property of waves, as opposed to particles. Here, we are dealing with a SHO, which would not seem like much of a wave just yet. I mean, it is not a traveling wave.[3] It can however be thought of as a

[2]Note that in this notation, f_i and $\tilde{f}$, if Fourier components, would include sinusoidal functions and multiplicative constants.

[3]If many SHOs are connected to each other, then we will have a traveling wave, as we will see later.

localized standing wave. We will deal more extensively with waves, near the end of this course.

As you can see, the superposition principle is a great property. It means that when an arbitrary form of a driving force $f(t)$ is given, then one can first decompose $f(t)$ to a sum (or, an integral) of convenient components, solve the problem for each component force, and then sum (or integrate) all solutions!

For instance, any "piecewise continuous" periodic function[4]can be written in a Fourier series.[5] Furthermore, any integrable function can be expressed as a Fourier integral. It then follows that if the response of a linear system to a force of a single sinusoid with an arbitrary angular frequency is known, then we know the response of the system to an arbitrary form of force.[6]

So, this is the reason why we solve the driven SHO problem with a sinusoidal driving force in the next section.

Before we do that, let us come back to the physics point of view of the superposition. We mentioned that the principle of superposition means that each solution $x_i(t)$ (or $\tilde{x}(\alpha, t)$) "remains intact" when they are put together. Put another way, this means that if the principle of superposition breaks down then when those solutions are put together something new happens. Indeed, if a non-linear interaction is included, then some new behaviors occur, as we will see when we discuss non-linear oscillations.

5.1.3 Driven SHO, a sinusoidal force

So, consider a sinusoidal driving force $F(t) = F_0 \cos(\omega t)$, *without loss of genarality*, as explained in the last section. Let $A = F_0/m$, and then the equation to solve becomes:

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x \;=\; A\cos(\omega t)$$

. . . Go to the complex world (I mean complex plane).
Turn this equation into

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x \;=\; A\exp(i\omega t)$$

[4] "Piecewise continuous" means that except at a finite number of points in any arbitrary finite interval, the function is continuous.

[5] Please familiarize yourself with Example 3.6 of the textbook. I won't quite require that you master it at this point, but you should understand the qualitative idea.

[6] Assuming that we can do the Fourier transform and the inverse-Fourier transform.

We solve for a *complex* $x(t)$ and then, at the end, can take the real part of it, as the actual EOM is the real part of the complex equation that we just set up. Assume

$$x(t) \;=\; C\exp(i\omega t)$$

Plugging this in, we get

$$C(-\omega^2 + 2\beta\omega i + \omega_0^2)\exp(i\omega t) \;=\; A\exp(i\omega t)$$

Since this equation should hold at any time, we get

$$C(-\omega^2 + 2\beta\omega i + \omega_0^2) \;=\; A$$

$$C \;=\; \frac{A}{\omega_0^2 - \omega^2 + 2\beta\omega i}$$

The magnitude of C is (call it D: $D \stackrel{def}{=} |C|$),

$$D \;=\; \frac{A}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\beta\omega)^2}}$$

As ω increases from 0, C goes into the 4th quadrant and then the 3rd quadrant, and so it makes sense to define its phase as $-\delta$. It is given by, from the above equation of C,

$$\delta \;=\; \tan^{-1}\left(\frac{2\beta\omega}{\omega_0^2 - \omega^2}\right)$$

Then,

$$x(t) \;=\; D\exp(-i\delta)\exp(i\omega t)$$

. . . Come back to the real world (I mean real axis).
Taking the real part, we get

$$x(t) \;=\; D\cos(\omega t - \delta)$$

with D and δ as given above. δ is the "**phase shift/lag**."

Before we go on further, let us note that the full solution is of the form

$$x_c(t) + x_p(t)$$

What we obtained just now is $x_p(t)$. We already know what $x_c(t)$ is from the previous lecture (the solution discussed in the damped SHO section). When a finite damping is present, $x_c(t)$ is always damped with the damping constant given by β or $\beta - \sqrt{\beta^2 - \omega_0^2}$ (over-damped). This means that if we wait for time $\gg 1/$damping constant, $x_c(t)$ is negligibly small. The solution that applies at such a large value of time is called the **steady state solution**. The solution that applies to the initial time when both $x_c(t)$ and $x_p(t)$ are appreciable is called the **transient solution.** Here, we will discuss the steady state solution only, that is, $x_p(t)$ only.

Let us analyze this phase lag/shift function $\delta(\omega)$ a bit.

$$\delta \approx 2\beta\omega/\omega_0^2 \qquad \text{when } \omega \approx 0$$

Positive slope. The slope increases as damping increases.

When $\omega \approx \omega_0$, $\frac{2\beta\omega}{\omega_0^2 - \omega^2} \approx \frac{2\beta\omega_0(1+\eta)}{-2\omega_0^2\eta} \approx -2\beta/(\omega_0\eta)$, where $\eta \stackrel{def}{=} (\omega - \omega_0)/\omega_0$. $\tan\delta \approx -2\beta/\eta$. Using $\tan\delta \approx \tan(\pi/2 + \tilde{\delta}) \approx -\cot(\tilde{\delta}) \approx -1/\tilde{\delta}$, where $\tilde{\delta} \stackrel{def}{=} \delta - \pi/2$, we get

$$\begin{aligned} \delta &\approx \frac{\pi}{2} + \frac{\omega_0\eta}{2\beta} \qquad \text{when } \omega \approx \omega_0 \\ &= \frac{\pi}{2} + \frac{\omega - \omega_0}{2\beta} \end{aligned}$$

$\delta = \pi/2$ at $\omega = \omega_0$, and the slope there goes to infinity as $\beta \to 0$. When $\beta \to 0$, $\delta = 0$ if $\omega < \omega_0$ and π if $\omega > \omega_0$, and so this infinite slope is reasonable.

OK, that is enough analysis for the phase lag/shift. You should compare Figure 3-16(b) of the textbook with this analysis.

How do we understand $\pi/2$ phase lag at $\omega = \omega_0$? If you are good at pumping a swing, convince yourself (or not??) that this is consistent, that is, you apply torque exactly the quarter cycle before the amplitude becomes maximum (or not??). Note that when $\delta = \pi/2$, the power is delivered to the system by the external force in the most optimum way, since $x = D\cos(\omega_0 t - \pi/2) = D\sin(\omega_0 t)$ and thus $v = D\omega_0\cos(\omega_0 t)$. The power $= Fv$, and F and v are exactly in phase, both behaving as $\cos(\omega_0 t)$, when $\omega = \omega_0$. Also, a homework problem should help you on this point.

Let us analyze the function $D(\omega)$ a bit as well. It is recognized immediately that if $\beta = 0$, then $D \to \infty$ at $\omega = \omega_0$. Thus, D in general is expected to have a peak structure around the natural frequency of the system. This is the so-called **resonance**. Near $\omega \approx 0$, $D \approx A/\omega_0$. The maximum of D is obtained by putting $dD/d\omega = 0$, which means $2(\omega^2 - \omega_0^2)2\omega + 8\beta^2\omega = 0$, which means $\omega = \sqrt{\omega_0^2 - 2\beta^2}$. This is the so-called **amplitude resonance frequency**, ω_R.

$$\omega_R = \sqrt{\omega_0^2 - 2\beta^2}$$

$\omega_R < \omega_1 = \sqrt{\omega_0^2 - \beta^2} < \omega_0$, where ω_1 is the oscillation frequency with under-damping.

If β is small, then $\omega_R \approx \omega_0$. Near the peak, consider the *intensity* profile: $D^2 \approx \frac{A^2}{(\omega^2-\omega_R^2)^2+(2\beta\omega_R)^2} \approx \frac{A^2}{(2\omega_R)^2[(\omega-\omega_R)^2+\beta^2]}$. This is the so-called Lorentzian line shape, centered at ω_R with the full width half maximum (FWHM), 2β. For this reason, it is customary to define the Q factor (quality factor):

$$Q \stackrel{def}{=} \frac{\omega_R}{2\beta}$$

As such, Q defines how sharp the resonance behavior is, as ω is swept. Poor quality oscillator with a large damping has a poor resonance characteristics. For instance, for a critically damped or a over-damped oscillator, the resonance characteristics will be very poor. High Q factor is required for precise measurements.

The resonance frequencies for the amplitude and the potential energy are identical, since the potential energy is proportional to the amplitude squared. However, the kinetic energy of the SHO is not necessarily maximized at the same frequency. As

$$x \propto \frac{1}{\sqrt{(\omega^2 - \omega_0^2)^2 + 4\omega^2\beta^2}} \cos(\omega t - \delta)$$

it follows that

$$\dot{x} \propto \frac{-\omega}{\sqrt{(\omega^2 - \omega_0^2)^2 + 4\omega^2\beta^2}} \sin(\omega t - \delta)$$

$$T \propto \dot{x}^2 \propto \frac{\omega^2}{(\omega^2 - \omega_0^2)^2 + 4\omega^2\beta^2} \sin^2(\omega t - \delta)$$

Let us use **the following *very important* result**

$$\frac{1}{T}\int_{t_0}^{t_0+T} dt \sin^2(\omega t - \delta) = \frac{1}{2} \qquad \text{where } T \stackrel{def}{=} 2\pi/\omega \text{ and } t_0 \text{ is any real number}$$

Then, the average kinetic energy $\langle T \rangle$ over the period T is given by

$$\langle T \rangle \propto \frac{\omega^2}{(\omega^2 - \omega_0^2)^2 + 4\omega^2\beta^2} \frac{1}{2}$$

Taking the ω-derivative, $d\langle T \rangle/d\omega \propto \omega[(\omega^2-\omega_0^2)^2+4\omega^2\beta^2] - \omega^2[(\omega^2-\omega_0^2)2\omega+4\omega\beta^2] = \omega(-\omega^4+\omega_0^4)$. So, $\langle T \rangle$ has a maximum at $\omega_E = \omega_0$.[7] be the **kinetic energy resonance frequency**, occurring exactly at the natural frequency of the system, different from the amplitude/potential-energy resonance frequency $= \omega_R = \sqrt{\omega_0^2 - 2\beta^2}$.

5.2 Examples 3.4 and 3.5

There is an exact analogy between the circuit and the mechanical system.

$$L\frac{dI}{dt} + \frac{Q}{C} + RI = \mathcal{E}$$

Noting that $I = \dot{Q}$, the circuit equation becomes

$$L\ddot{Q} + R\dot{Q} + \frac{1}{C}Q = \mathcal{E}$$

The following correspondence can then be noted.

Circuit	Mechanical
Q	x
L	m
R	b
$1/C$	k
$\mathcal{E}$	F

Example 3.4 describes a hanging mass-spring system with an equivalent LC circuit. (Left for reading.)

[7] Here, we just showed that $\langle T \rangle$ has a unique extremum at $\omega = \omega_E$ for $\omega >= 0$. How do we know it is maximum? If you are not sure, do NOT take the 2nd derivative. That would be too complicated. Rather, note that $\langle T \rangle \to 0$ when $\omega \to 0$ or ∞, while $\langle T \rangle$ = finite at $\omega_E = \omega_0$. Since ω_E is the only extremum point between $\omega = 0$ and $\omega = \infty$, it must be a maximum.

Example 3.5 is an RLC resonant circuit, driven by an emf $\mathcal{E} = E_0 \cos \omega t$ (correct textbook). The problem is to find the resonance frequency to maximize V_L. The voltage $V_L = L\ddot{Q}$. So, this is equivalent to maximizing the acceleration.

$$\begin{aligned} \omega_0^2 &= k/m \to 1/(LC) \\ \beta &= b/(2m) \to R/(2L) \\ A &= F_0/m \to E_0/L \end{aligned}$$

$$\begin{aligned} Q &= \frac{E_0}{\sqrt{(L\omega^2 - 1/C)^2 + R^2\omega^2}} \cos(\omega t - \delta) \\ I &= \frac{-E_0 \sin(\omega t - \delta)}{\sqrt{(L\omega - 1/(C\omega))^2 + R^2}} \\ V_L &= LdI/dt = \frac{-E_0 \omega \cos(\omega t - \delta)}{\sqrt{(L\omega - 1/(C\omega))^2 + R^2}} \\ &\stackrel{def}{=} V(\omega) \cos(\omega t - \delta) \end{aligned}$$

To maximize $V(\omega)$, note that $d|V(\omega)|^2/d\omega \propto 2\omega((L\omega - 1/(C\omega))^2 + R^2) - \omega^2 2(L\omega - 1/(C\omega))(L + 1/(C\omega^2)) \propto (LC\omega^2 - 1)^2 + R^2C^2\omega^2 - (LC\omega^2 - 1)(LC\omega^2 + 1) \propto (R^2C^2 - 2LC)\omega^2 + 2$. Putting this to be 0, we get

$$\omega_{max} = \frac{1}{\sqrt{LC - \frac{R^2C^2}{2}}}$$

This is the frequency at which the acceleration is maximized. It is different from $\omega_E = \omega_0 = 1/\sqrt{LC}$ and the amplitude resonant frequency $\omega_R = \sqrt{\omega_0^2 - 2\beta^2} = \sqrt{1/LC - R^2/(2L^2)}$.

5.3 Non-linear oscillations

5.3.1 Phase space, 1st order ODE

Newton's equation is a 2nd order ODE in the real space. However, in the phase space it is a 1st order ODE. Why? Because the equation

$$m\ddot{\vec{r}} = \vec{F}(\vec{r}, \dot{\vec{r}}, t)$$

becomes a set of equations

$$\begin{aligned} \dot{\vec{v}} &= \vec{F}(\vec{r}, \vec{v}, t)/m \\ \dot{\vec{r}} &= \vec{v} \end{aligned}$$

In this new representation, Newton's law becomes a first order equation in the $(\vec{r}, \vec{v})$ space. Generalization of this new representation to a multi-particle case is straightforward: just take $\vec{r}$ to include all position coordinates of all particles, and $\vec{v}$ to include all velocity components of all particles.

With the language of the phase space, and the above set of equations in the phase space, what we discussed about Newtonian determinism becomes very lucid. It means that given an initial condition, a point is specified in the phase space. Then, the time evolution of that point is dictated by Newton's 2nd law as a mapping from the initial point to the next point at the next time (dt) and then at a later time ($2dt$) and so on and so forth. In this sense, **classical dynamics can be thought of as a mapping of a point in phase space**.

The numerical integration of Newton's EOM usually sets up the equation in phase space, as shown above, and do a single integration, using, e.g., the Runge-Kutta method.

5.3.2 Linear and non-linear systems

In a linear system, if two slightly different inputs go in, then "slightly different outputs" come out. Here, slightly different outputs mean the outputs whose difference is linearly proportional to the initial difference.

This sounds reasonable. But, let us be a bit more analytical.

Consider dividing up time into many tiny steps of Δt, which is much smaller than the experimental resolution. Newton's EOM can then be written as a succession of discrete transformations:

$$\vec{P}_{n+1} = \vec{P}_n + (\vec{v}_n, \vec{F}(\vec{P}_n, t_n))\Delta t$$

$\vec{P}$ is a point in phase space, $(\vec{r}, \vec{v})$, and $n = 0, 1, 2, \cdots$. $\Delta t = t/N$, where N is a large natural number, and $t_n = n\Delta t$. Taking this recurrence relation, a computer

can use this equation to find the solution at any specified time t in the future or in the past. Note that Δt here is basically dt, but I am using Δt for notational convenience.

The function $\vec{F}(\vec{P},t)$ can be written as

$$\vec{F} \;=\; \vec{f}(t) + \vec{g}(\vec{P},t)$$

where $\vec{f}(t)$ is a term that is purely dependent on time (like the driving force term of a driven SHO), and $\vec{g}$ is a function that explicitly depends on dynamical variables ($\vec{P}$) with a possible dependence on t as well.[8]

> A linear system is defined as a system where $\vec{g}(\vec{P},t)$ is a linear function of $\vec{P}$.

The SHO (damped or driven) we discussed so far is a linear system. In 1D, $\vec{g}(\vec{P}) = -kx - bv$. This is a linear function on a phase space vector $\binom{x}{v}$.

Now, let us look at the mathematics of the transformation from $\vec{P}_n$ to $\vec{P}_{n+1}$ for a linear system. We start from the initial condition $\vec{P}_0$.

$$\vec{P}_1 \;=\; \vec{P}_0 + (\vec{v}_0, \vec{F}(\vec{P}_0,0))\Delta t$$

Say, we have a slightly different initial condition $\vec{P}_0{}' \;=\; \vec{P}_0 + \varepsilon\vec{d}$, with $\vec{d} \;=\; (\vec{r}_d, \vec{v}_d)$, and ε is a very small number. What is the result of the transformation for $\vec{P}_0{}'$?

$$\begin{aligned}
\vec{P}_1{}' \;&=\; \vec{P}_0 + \varepsilon\vec{d} + (\vec{v}_0 + \varepsilon\vec{v}_d, \vec{F}(\vec{P}_0 + \varepsilon\vec{d},0))\Delta t \\
&=\; \vec{P}_0 + \varepsilon\vec{d} + (\vec{v}_0 + \varepsilon\vec{v}_d, \vec{F}(\vec{P}_0) + \varepsilon\vec{g}(\vec{d},0))\Delta t \qquad \vec{F}(\vec{P}_0{}',0)) \;=\; \vec{F}(\vec{P}_0,0) + \varepsilon\vec{g}(\vec{d},0) \\
&=\; \vec{P}_1 + \varepsilon\left[\vec{d} + (\vec{v}_d, \vec{g}(\vec{d},0))\Delta t\right]
\end{aligned}$$

Now, we will define $\vec{d}_0 \overset{def}{=} \vec{d}$, and define a recurrence relation for the *"difference vector* (up to a multiplicative constant ε)" $\vec{d}_n$, as follows.

$$\vec{d}_{n+1} \;=\; \vec{d}_n + (\vec{v}_n, \vec{g}(\vec{d}_n, t_n))\Delta t$$

[8] An example of this is a person who pumps a swing by simply sitting up and down to change a parameter (l) of the system. This mechanism of pumping a swing is called the "parametric resonance" and it is different from the resonance that we discussed in the previous section. Please look it up if you are interested.

Note that the difference vector evolves in time only by $\vec{g}$. This is reasonable since the effect of $\vec{f}$ is does not depend on $\vec{P}$.

So, what we have is the following.

$$\begin{aligned}
\vec{P}_0{}' &= \vec{P}_0 + \varepsilon\vec{d}_0 \\
\vec{P}_1{}' &= \vec{P}_1 + \varepsilon\vec{d}_1 \\
\vec{P}_2{}' &= \vec{P}_2 + \varepsilon\vec{d}_2 \qquad \text{Repeat what we did for } n = 0 \to 1. \\
&\cdots \\
\vec{P}_n{}' &= \vec{P}_n + \varepsilon\vec{d}_n \qquad \text{for any } n.
\end{aligned}$$

What is important to realize is that in this series, each step is *exact*. If we had a non-linear system, then a step might be approximately of the same form as what we wrote down here, but a cumulative effect of all steps may be completely non-linear in ε. What we showed here amounts to saying that such cannot be the case for a linear system, due to the exact linear nature of each step that is investigated above.

So, we have proved the following important fact.

> For a linear system, a small difference in the initial condition produces a difference in $\vec{P}$, which is only linearly proportional to the initial difference.

This is what we mean when we casually say "a small difference in inputs means a small difference in outputs." For instance, two balls rolled in slightly different directions will separate from each other with a very large difference in the positions as time becomes very large. But, the point is that such a difference will still be linear in the initial difference, and so it is "easy for us to understand it." If we consider a bound motion, such as two identical SHOs with slightly different initial conditions, such as slightly different phases (positions), then it is obvious that at all times the difference in positions will remain very small ($x_1(t) = A\cos(\omega t)$ and $x_2(t) = A\cos(\omega t + \varepsilon)$: $x_1 - x_2 \approx \varepsilon A\sin(\omega t) = O(\varepsilon)$ at *all times*). Why is this? Well, we can look at the above equation that we derived: $\vec{P}_n{}' = \vec{P}_n + \varepsilon\vec{d}_n$. If ε is vanishingly small, and if we have a bound motion (all $\vec{P}, \vec{P}', \vec{d}$ vectors are finite in size), then we can immediately see the following property.

> Consider two linear systems undergoing bound motions. Assume that the two systems are identical to each other, except for a vanishingly small difference in

the initial conditions. Then, at any time, the difference between the two motions is vanishingly small, i.e. they are represented by two nearly identical points in the phase space at any given time.

The very defining characteristics of **chaos** is that this property breaks down. A seemingly negligible difference in the initial conditions between two bound state motions causes completely different results at a later time! Two almost identical inputs go in, and completely different results can come out, since the results exponentially diverge from each other, not linearly. This **sensitive dependence on initial conditions** is generally called chaos, while a more quantitative definition will be given later.

Some non-linear systems show chaotic behaviors. Non-linear problems are generally hard to solve by hand, and chaotic problems are, by definition, impossible to solve by hand. So the computer plays an important role in this topic.

5.3.3 Non-linear oscillations

For a real potential function that is encountered,[9] the deviation from the quadratic behavior of a SHO is a rule than an exception. Suppose that there is a potential minimum at $x = 0$. In general, then, one should write,

$$U(x) \;=\; U_0 + \frac{1}{2}m\omega_0^2 x^2 + \frac{1}{3}m\alpha x^3 + \frac{1}{4}m\beta x^4 + \cdots$$

$$F(x) \;=\; -m\omega_0^2 x - m\alpha x^2 - m\beta x^3 + \cdots$$

When the α term and higher are not negligible, we have "non-linear oscillations." $F(x)$ is clearly non-linear with those terms included: for example, $(ax_1 + bx_2)^2 \neq ax_1^2 + bx_2^2$!

Every oscillation in nature is non-linear, as a rule. For instance, the reason why things tend to expand when they get hot is because of the non-linear oscillations inside materials (we can see the qualitative reason using classical mechanics, as we will do shortly).

It is reasonable to do a perturbation expansion on non-linear terms to get their leading order correction. This is done in Landau § 28, with the amplitude of the

[9] E.g. the Lennard- Jones potential.

small oscillation as the perturbation parameter. Here, we will not pause to derive[10] the following result. However, it is of interest to understand the qualitative physics of the result. In the following, the initial phase (θ_0) of the motion is taken to be 0 (by shifting the origin of time).

$$x(t) \approx A\cos(\omega t) - \frac{\alpha A^2}{2\omega_0^2} + \frac{\alpha A^2}{6\omega_0^2}\cos(2\omega t)$$

$$\omega \approx \omega_0 + \left(\frac{3\beta}{8\omega_0} - \frac{5\alpha^2}{12\omega_0^3}\right) A^2$$

Characteristics of non-linear oscillators

Here are some essential notable things about this solution.

- **New effective equilibrium position** The time averaged value of x is not zero any more, but $-\frac{\alpha A^2}{2\omega_0^2}$. And the offset is amplitude dependent!
- **Overtones/Higher-harmonics** Notice that $x(t)$ has frequency 2ω. At higher orders of perturbation, further multiples of ω will appear.
- **Frequency softening/hardening** The angular frequency ω is now different from ω_0. Assuming $\alpha, \beta > 0$, we see that β increases the frequency (hardening), and α decreases it (softening).

How do we understand these characteristics?

- **New effective equilibrium position** Assuming $\alpha > 0$, it is offset to the negative side. Why? Because that is where the leading order potential energy correction $\propto \alpha x^3$ is negative. So, while $x = 0$ is still the equilibrium point in a mathematical sense, it is meaningless to call that the equilibrium position. The more meaningful equilibrium position is the time average value of x, which is the observable of physics. In terms of this effective equilibrium position, the particle is no longer at 0. Instead, the particle is shifted towards a more attractive part, which is the $x < 0$ region, and thus the effectively speaking the

[10] Interested readers should read Landau § 28. The perturbation method is slightly more complicated than the usual kind.

mean position of x shifts more and more towards left as the amplitude of the oscillation A becomes larger and larger. This is precisely why when materials are heated up and all those molecules inside them make more vibrations they generally expand (because the potential between molecules is hard – steep slope – on short distance and soft – gentle slope – on large distance).

- **Overtones/Higher-harmonics** That the overtones would occur in non-linear oscillators is not that surprising from the perturbation point of view. The perturbation terms like $-m\alpha x^2$ or higher terms will give the driving force of $\cos^2(\omega t)$ and higher powers. They create sinusoidal functions with angular frequency 2ω and higher multiples of ω. In non-linear optics, this type of phenomenon is extremely important. Typically a non-linear crystal is illuminated with intense laser light. Due to the non-linear effect, higher frequency light comes out. Light of frequency as high as $O(100)$ times that of the original light is generated by repeating such process.

- **Frequency softening/hardening** As non-linear terms in the potential modify the overall shape of the potential, it is not surprising that the natural frequency changes as well. The above result shows that, for small perturbation, the cubic term softens the frequency, i.e. decreases the frequency, while the quartic term hardens the frequency.