II. Perturbation Theory.

Most problems in nonlinear applied mathematics, at least most of those that can be solved, involve some sort of small or large parameter. Usually this is justified from the physical application. In perturbation theory we try to approximate a solution in terms of a sum of simpler functions which are successively less and less important. For example, we can consider the algebraic problem

\[ x^3 - \varepsilon x - 1 = 0, \quad (1) \]

to be solved for \( x \). One approach is to use the implicit function theorem. At \( \varepsilon = 0 \) there is the unique real solution \( x = 1 \). Letting \( f(x, \varepsilon) = x^3 - \varepsilon x - 1 = 0 \), we have

\[ f_x(1, 0) = 3x^2 - \varepsilon \big|_{(1, 0)} = 3. \]

Therefore for small \( \varepsilon \) there is a unique solution \( x(\varepsilon) \) such that \( x(0) = 1 \). This solution is differentiable with respect to \( \varepsilon \) near \( \varepsilon = 0 \). Differentiating with respect to \( \varepsilon \) gives

\[ (3x^2 - \varepsilon) \frac{dx}{d\varepsilon} - x = 0 \]

so \( \frac{dx}{d\varepsilon} \big|_{\varepsilon=0} = \frac{1}{3} \). Further, solving for \( \frac{dx}{d\varepsilon} \) we see that \( \frac{d^2x}{d\varepsilon^2} \) exists for \( \varepsilon \) small, and so by Taylor’s theorem,

\[ x(\varepsilon) = 1 + \frac{1}{3} \varepsilon + O(\varepsilon^2) \]

as \( \varepsilon \to 0 \).

But this approach gets tedious when looking for higher order terms. The usual technique is to assume a form

\[ x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots \]

and substitute this into (1). This gives

\[ (x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots)^3 - \varepsilon (x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots) - 1 = 0. \]
We then collect different powers of $\varepsilon$ and equate the coefficients. We have

$$x_0^3 - 1 + (3x_0^2 x_1 - x_0) \varepsilon + (3x_0^2 x_2 + 3x_0 x_1^2 - x_1) \varepsilon^2 + \cdots = 0.$$  

This yields the successive equations

$$x_0^3 - 1 = 0$$
$$3x_0^2 x_1 - x_0 = 0$$
$$3x_0^2 x_2 + 3x_0 x_1^2 - x_1 = 0$$

and so forth, and solving these one at a time we get $x_0 = 1$, $x_1 = \frac{1}{3}$, $x_2 = 0$, etc. Therefore we have the expansion

$$x = 1 + \frac{1}{3} \varepsilon + O (\varepsilon^3)$$

for small $\varepsilon$. This can be made rigorous for this particular example, but we will not take time to do this here.

Not all algebraic problems are this easy. Consider

$$x^3 - x^2 + \varepsilon = 0.$$  

Trying $x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots$, we get

$$x_0^3 - x_0^2 + \varepsilon (3x_0^2 x_1 - 2x_0 x_1 + 1) + \varepsilon^2 (3x_0^2 x_2 + 3x_1^2 x_0 - 2x_0 x_2) + O (\varepsilon^3),$$

so

$$x_0^3 - x_0^2 = 0$$
$$3x_0^2 x_1 - 2x_0 x_1 + 1 = 0$$

and so forth. Either $x_0 = 1$ or $x_0 = 0$. If $x_0 = 1$ then we get $x_1 = -1$ and can continue with $x_2$, etc. If $x_0 = 0$ then there is no solution to the second equation, for $x_1$.

We have to notice that near $x = 0$, $x^3 \ll x^2$ (this means: $\frac{x^3}{x^2} \to 0$ as $x \to 0$) and so we expect the solution to be approximately $\pm \sqrt[3]{\varepsilon}$. Therefore we try a solution

$$x = x_1 \sqrt[3]{\varepsilon} + x_2 \varepsilon + x_3 \varepsilon^{3/2} + \cdots$$
This will work out and we will be able to find the coefficients. Again this can be made rigorous.

We now turn to differential equations. The context will be perturbations of equations which have periodic solutions. We recall the following two theorems, which can be found in Coddington and Levinson:

Theorem p1: Consider a system

\[ x' = f(t, x, \mu) \tag{2} \]

where \( f \) is periodic in \( t \) of period \( T \). Assume also (for simplicity) that \( f_x \) and \( f_\mu \) are continuous for all \((t, x, \mu)\). Suppose that when \( \mu = 0 \), (2) has a periodic solution \( p \) of period \( T \). Suppose that the first variation

\[ y' = f_x(t, p(t), 0) y \]

has no non-trivial solution of period \( T \). Then for small \( |\mu| \) the equation (2) has a solution \( q = q(t, \mu) \) which is periodic in \( t \) of period \( T \), continuous in \((t, \mu)\), and with \( q(t, 0) = p(t) \). There is only one such solution for each small \( \mu \).

Before stating the second theorem, we recall that in the autonomous case, where \( f \) does not depend on \( t \), the hypotheses cannot be satisfied if \( p \) is a non-constant periodic solution. If \( p(t) \) is a periodic solution of

\[ x' = f(x, 0) , \]

then by differentiating the equation \( p' (t) = f(p(t), 0) \) with respect to \( t \) we see that \( p'(t) \) is a periodic solution of the linearized equation

\[ y' = f_x(p(t), 0) y . \]

Theorem p2: Under the same differentiability hypotheses on \( f \), suppose that \( p(t) \) is a non-constant periodic solution of

\[ x' = f(x, 0) , \]

and suppose that 1 is a simple characteristic root (Floquet multiplier) for the periodic linear system

\[ y' = f_x(p(t), 0) y . \]
Then for small $|\mu|$ the system
\[
x' = f(x, \mu)
\]
has a solution $q(t, \mu)$ of period $T(\mu)$. Both $q$ and $T$ are continuous in $\mu$ for sufficiently small $|\mu|$, with $q(t, 0) = p(t)$ and $T(0) = T$.

Our goal in this section is to study cases where these theorems do not apply. Here is a simple example (with $\varepsilon$ as the parameter instead of $\mu$):
\[
x' = \varepsilon \sin t \ x.
\]
The unperturbed equation is
\[
x' = 0,
\]
and so any constant is a periodic solution. The linearized equation is the same, and so neither theorem applies. We can get a unique periodic solution by adding an initial condition: say
\[
x(0) = 1.
\]
We can solve the complete equation exactly:
\[
x(t, \varepsilon) = e^{-\varepsilon \cos t}.
\]
Thus there is a periodic solution for every $\varepsilon$, and $\lim_{\varepsilon \to 0} x(t, \varepsilon) = 1$ uniformly on the entire line $-\infty < t < \infty$.

Now consider the example
\[
x' = \varepsilon \sin^2 t \ x.
\]
This is still a periodic equation. The solution with $x(0) = 1$ is
\[
x(t, \varepsilon) = e^{c(t - \frac{\sin 2t}{4})}.
\]
which is not periodic. In fact, it is unbounded, and so it is not true that $\lim_{\varepsilon \to 0} x(t, \varepsilon) = 1$ uniformly on the entire line $-\infty < t < \infty$.

It is true, however, that $\lim_{\varepsilon \to 0} x(t, \varepsilon) = 1$ uniformly on any compact interval $[-M, M]$. But this is not saying much. We recall that solutions of the initial value problem $x' = \varepsilon \sin^2 t \ x, \ x(0) = 1$ are continuous with respect to $\varepsilon$. To be more precise, the function $x(t, \varepsilon)$ which solves this initial value problem is continuous in the pair $(t, \varepsilon)$ for each $(t, \varepsilon)$ where the solution exits. It is not hard to show that the solution exists for all $(\tau, \varepsilon)$. So from advanced calculus we know that $x(t, \varepsilon)$ is
uniformly continuous on any compact set. In particular, it is uniformly continuous on
\[ \{(t, \varepsilon) | 0 \leq \varepsilon \leq 1, -M \leq t \leq M \}, \]
and so the statement that \( \lim_{\varepsilon \to 0} x(t, \varepsilon) = x(t, 0) \) is true uniformly on \(|x| \leq M\). No particular information about the exact solution is required for this conclusion. But knowing the exact solution we can say something stronger. Consider any interval \(-M(\varepsilon) \leq t \leq M(\varepsilon)\) with the property that \( \lim_{\varepsilon \to 0+, \varepsilon M(\varepsilon) = 0} \). We see from (3) that
\[ \lim_{\varepsilon \to 0^+} \sup \{ |x(t, \varepsilon) - 1| \mid -M(\varepsilon) \leq t \leq M(\varepsilon) \} = 0. \]
For example, we could have \( M(\varepsilon) = |\log \varepsilon| \). (Thus, \( M(\varepsilon) \) can be unbounded as \( \varepsilon \to 0 \).) In this case we would say that 1 approximates \( x(t, \varepsilon) \) uniformly “on a scale of \( O(|\log \varepsilon|) \). This is more than we get from the basic fact that solutions are continuous with respect to \( \varepsilon \). Note, however, that we do not get uniform approximation on a scale of \( O\left(\frac{1}{\varepsilon}\right) \).

Later we will consider the problem of finding approximate solutions on some set in more detail.

But first we consider another example, this time autonomous. Theorem p2 will not apply. The equation is van der Pol’s equation, written as
\[ x'' + \varepsilon (x^2 - 1) x' + x = 0. \]
(4)
This is an important example which we will consider at some length.

The unperturbed equation is
\[ x'' + x = 0. \]
Obviously this has many periodic solutions. You should check that the hypotheses of Theorem p2 do not apply. To within a translation in \( t \) they can all be written as
\[ x = A \cos t. \]
Any \( A \) will do. Yet, we know that for positive \( \varepsilon \), (4) has a unique periodic solution, say \( x(t, \varepsilon) \). Presumably, as \( \varepsilon \to 0 \), this solution tends to one of the periodic solution when \( \varepsilon = 0 \). The problem is, which one?

There are many approaches to this problem. Let’s start with the “naive expansion” approach. This means that we look for a solution
\[ x(t, \varepsilon) = x_0(t) + \varepsilon x_1(t) + \cdots. \]
Substituting into the differential equation we have

\[
x''_0 + \varepsilon x''_1 + \cdots + \varepsilon ( (x_0(t) + \varepsilon x_1(t) + \cdots)^2 - 1 ) (x'_0(t) + \varepsilon x'_1(t) + \cdots) + x_0(t) + \varepsilon x_1(t) + \cdots = 0.
\]

Equating coefficients of different powers of \( \varepsilon \), we get

\[
x''_0 + x_0 = 0
\]
\[
x''_1 + (x''_0 - 1) x'_1 + x_1 = 0,
\]

and so forth. (This is all we need to get an interesting conclusion.) We can write the general \( x_0 \) as

\[x_0 = A \cos (t - t_0),\]

but \( A \) is still to be determined. We can choose the initial condition so that \( t_0 = 0 \). We then have

\[x''_1 + x_1 = (1 - A^2 \cos^2 (t)) (-A \sin (t)) = A^3 \sin t \cos^2 t - A \sin t.
\]

This is a linear equation which can be solved by the variation of parameters method. Being lazy, we use Maple and get the general solution as \( x_1(t) = -\frac{1}{8} A^3 \sin t \cos^2 t - \frac{1}{8} (\cos t) A^3 t + \frac{1}{2} (\cos t) At + C_1 \cos t + C_2 \sin t \).

We are looking for a periodic solution. Recall that we have not specified \( A \). Two of the terms for \( x_1 \) have factors of \( t \), which are not periodic, and which in fact are unbounded. If we are to find a periodic solution we must choose \( A \) so the sum of these terms vanishes. (This is called a “non-resonance” condition, or a condition to “eliminate secular terms”.) This gives \(-\frac{A^3}{8} + \frac{A}{2} = 0\), or \( A^2 = 4 \). Therefore, we conjecture that the amplitude of \( x_0 \) should be \( A = 2 \). However:

**This is not a rigorous proof!**

Nevertheless, we may be impressed that this naive approximation method seems to give an answer. So before turning to rigorous methods, we consider another example: We will now try to apply the method to the following equation:

\[x'' + x + \varepsilon x^3 = 0, \quad (5)\]

looking for periodic solutions. Once again we try

\[x = x_0(t) + \varepsilon x_1(t) + \cdots,
\]
and we find the equations

\[ x_0'' + x_0 = 0 \]
\[ x_1'' + x_1 = -x_0^3. \]

We set \( x_0(t) = A \cos t \) and solve the \( x_1 \) equation:

\[ x_1'' + x_1 = -A^3 \cos^3 t \quad (6) \]

The exact solution is: \( x_1(t) = \frac{1}{8} A^3 \cos^3 t - \frac{3}{8} A^3 \cos t - \frac{3}{8} A^3 (\sin t) t + C_1 \cos t + C_2 \sin t \)

Now we find only one resonant term, the second, and to get rid of it, the only possibility is that \( A = 0 \).

So we ask: Does that mean that our equation has no periodic solutions except possibly some with small amplitude, order \( \varepsilon \)?

We know this is not true. There is an energy function for (5), namely \( \frac{1}{2} x'^2 + \frac{1}{2} x^2 + \frac{\varepsilon}{4} x^4 \), and all of the curves \( E = \text{constant} \) give periodic solutions. These can be arbitrarily large in amplitude, for any \( \varepsilon \geq 0 \). So why don’t we find these with our asymptotics?

The answer is seen by looking carefully at what is implied by the equation

\[ x(t, \varepsilon) = x_0(t) + \varepsilon x_1(t) + O(\varepsilon^2). \]

Suppose that all of the \( x_k(t) \) are periodic with some period \( T \). Then we are expecting that as \( \varepsilon \to 0 \), the periodic solution tends uniformly in time (since we only have to look at a fixed interval \([0, T]\)) to a limiting nontrivial periodic solution \( x_0(t) \). This does not allow for the possibility that the period of \( x(t) \) will probably not be the same as the period of \( x_0(t) \), and therefore, for any small \( \varepsilon \), the solution will eventually drift away from \( x_0(t) \), even if the orbits of \( x_0(t) \) and \( x(t, \varepsilon) \) are very close. In fact, this could, and does, happen for van der Pol’s equation also. So our derivation of the limiting value of \( A = 2 \) in van der Pol’s equation is very suspicious indeed. We now consider an important method for resolving this difficulty, called the

**Method of Averaging.**

We will now give still another approach to van der Pol’s equation, and this one we will make rigorous.
This material appears in GH, but here I will also use the Grimshaw reference given at the beginning of the course. We will start by reconsidering one of our first order examples, namely

\[ x' = \varepsilon f(x, t, \varepsilon) = \varepsilon \sin^2 t \ x. \]
\[ x(0) = 1. \]

(We introduce \( \varepsilon \) into the function \( f \) for greater generality later on, but in this example, \( f \) does not depend on \( \varepsilon \).)

First we look on an \( O(1) \) time scale. Consider a fixed interval \([0, T]\). From the basic theorem about continuity of solutions with respect to parameters, it follows that

\[ \lim_{\varepsilon \to 0} x(t, \varepsilon) = 1. \]

uniformly on \([0, T]\). However we saw before that \( x = 1 \) is a valid uniform approximation over a larger time scale, such as \( 0 \left( \frac{1}{\sqrt{\varepsilon}} \right) \). For example, suppose that \( \varepsilon = 10^{-4} \). The approximation is valid over \([0, 100]\). We notice that in this interval \( x' \) has about seven oscillations. The method of averaging is based on an assumption that the oscillations will somehow average out, and allow us to find an approximation valid on even a longer time interval.

The way we average is to let

\[ \mathcal{F}(x, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} f(x, t, \varepsilon) \, dt. \]

In this case the result is simple:

\[ \mathcal{F}(x, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} x \sin^2 t \, dt = \frac{1}{2\pi} x \left( \frac{t}{2} - \frac{\sin 2t}{4} \right) \bigg|_0^{2\pi} = \frac{x}{2}. \]

The so called “averaged” equation is

\[ y' = \varepsilon \mathcal{F}(x, 0) = \varepsilon \frac{x}{2}. \]
with solution \( y(t) = e^{\varepsilon t^2} \). We then have:

\[
|x(t, \varepsilon) - y(t, \varepsilon)| = e^{\varepsilon t^2} \left( 1 - e^{-\varepsilon \sin 2t} \right)
\]

For a fixed \( \varepsilon \), this is unbounded as \( t \to \infty \), so we do not have uniform convergence on \([0, \infty)\). However, if \( t = \frac{K}{\varepsilon} \) for some fixed \( K \), then

\[
|x(t, \varepsilon) - y(t, \varepsilon)| = e^{\frac{K}{\varepsilon^2}} \left( 1 - e^{-\varepsilon \sin \frac{K}{4}} \right)
\]

which does tend to zero as \( \varepsilon \to 0 \). For \( 0 \leq t \leq \frac{K}{\varepsilon} \) we have

\[
|x(t, \varepsilon) - y(t, \varepsilon)| = e^{\frac{K}{\varepsilon^2}} \left( 1 - e^{-\varepsilon \sin \frac{K}{4}} \right) \leq e^{\frac{K}{\varepsilon^2}} (e^{\frac{K}{4}} - 1) .
\]

(You need to compare \( 1 - e^{-\varepsilon^2/4} \) and \( e^{\frac{K}{4}} - 1 \) to verify this.) Hence

\[
\lim_{\varepsilon \to 0} |x(t, \varepsilon) - y(t, \varepsilon)| = 0
\]

uniformly on \([0, \frac{K}{\varepsilon}]\), and so \( y \) approximates \( x \) on a time scale of \( O\left(\frac{1}{\varepsilon}\right)\). This is true not just for a periodic solution (there aren’t any for this equation) but for any solution. We will see in higher order examples that this is a very useful method.

In the method of averaging as applied to nonlinear oscillators such as van der Pol’s equation or Duffing’s equation, we consider a system in polar coordinates of the form

\[
\begin{align*}
r' &= \varepsilon F(r, \theta, \varepsilon) \\
\theta' &= \omega(t) + \varepsilon G(r, \theta, \varepsilon)
\end{align*}
\]

where \( \omega(t) \) is bounded away from zero and \( F \) and \( G \) are periodic with period \( 2\pi \) in \( \theta \). In our applications we will have \( \omega(t) = -1 \) for all \( t \). This means that for small \( \rho \), \( r \) changes slowly compared with \( \theta \). Therefore we expect that on some time scale we may be able to approximate the solution by averaging the \( r \) equation with respect to \( \theta \). We let

\[
\bar{F}(r, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} F(r, \theta, \varepsilon) \, d\theta
\]

and consider the equation

\[
R' = \varepsilon \bar{F}(R, 0) .
\]
Since $r' \to 0$ as $\varepsilon \to 0$, and also $R' \to 0$, we expect that if $R(0) = r(0)$, then $R$ and $r$ will both be close to $r(0)$ over an $O(1)$ time scale. This is just the statement that the solutions are uniformly continuous in $\varepsilon$ over a compact interval $[0, T]$, where $T$ is constant. It turns out, however, that $r$ is close to $R$ over the much longer interval $O(\frac{1}{\varepsilon})$. In such an interval, $r$ can change significantly and so by studying the single equation for $R$, we get interesting information about just how $r$ changes. The complete solution $(r, \theta)$ consists of the slow but significant change in $r$ on the $O(\frac{1}{\varepsilon})$ time scale and the relatively fast oscillations due to changes of $\theta$ on the $O(1)$ time scale.

Grimshaw uses rather different notation from what I have used. GH presents similar material, (chapter 4) with still different notation. Because I am mainly interested in the two applications we have already studied, I will stick with $r$ and $R$.

We consider first the autonomous case, with a second order equation of the form

$$u'' + u = \varepsilon f(u, u', \varepsilon).$$

(9)

Note that both van der Pol’s equation and Duffing’s equation in the form above can be written like this. We then let

$$x = u, y = u'$$

and write a system

$$x' = y$$

$$y' = -x + \varepsilon f(x, y, \varepsilon).$$

(10)

Switching to polar coordinates we write

$$x = r \cos \theta, y = r \sin \theta$$

so that

$$x' = r' \cos \theta - (r \sin \theta) \theta' = r \sin \theta$$

$$y' = r' \sin \theta + (r \cos \theta) \theta' = -r \cos \theta + \varepsilon f(r \cos \theta, r \sin \theta, \varepsilon).$$

(11)
We can easily solve these, by multiplying the equations by $\cos \theta$ and $\sin \theta$ and adding and subtracting in the right way, and we get
\[
\begin{align*}
r' &= \varepsilon \sin \theta f (r \cos \theta, r \sin \theta, \varepsilon) = \varepsilon F (r, \theta, \varepsilon) \\
\theta' &= -1 + \frac{\varepsilon}{r} \cos \theta f (r \cos \theta, r \sin \theta, \varepsilon) = -1 + \varepsilon G (r, \theta, \varepsilon).
\end{align*}
\]

Notice that both $F$ and $G$ are periodic in $\theta$ with period $2\pi$. We let
\[
F (r, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} F (r, \tau, \varepsilon) d\theta.
\]

We then use (7) and (8) to obtain the “averaged equation”
\[
R' = \varepsilon F (R, 0).
\]

Note that we have evaluated $F$ at $\varepsilon = 0$, but the equation for $R$ does have $\varepsilon$ on the right hand side.

We will state and prove a theorem concerning the relation between $R(t)$ and $r(t)$, but first let us try applying this to the same Duffing equation as before:
\[
x'' + x + \varepsilon x^3 = 0.
\]

Recall that all solutions are periodic. So the problem here is not to show that there are periodic solutions. However the period of the solution depends on its amplitude. We can ask, for example, whether the period increases or decreases with amplitude, and by how much for a given small $\varepsilon$. The energy function
\[
E (t) = \frac{x'^2}{2} + \frac{x^2}{2} + \frac{x^4}{4}
\]

can be used to determine this but we can also use averaging, at least for small $\varepsilon$.

The equations for $r$ and $\theta$ become:
\[
\begin{align*}
r' &= -\varepsilon \sin \theta r^3 \cos^3 \theta \\
\theta' &= -1 - \varepsilon r^2 \cos^4 \theta.
\end{align*}
\]

11
Therefore
\[ F(r, \varepsilon) = -\frac{1}{2\pi} \int_0^{2\pi} \sin \theta \cos^3 \theta \, d\theta = 0 \]
and so the averaged \( R \) equation is
\[ R' = 0. \]

Therefore, \( R = A = \text{constant} \).

Usually the method stops with the equation for \( R \), but let us proceed to study \( \theta \).

If we let \( \psi(t) = \theta(t) + t \) we get
\[ \psi' = -\frac{\varepsilon}{r} \cos (\psi - t) r^3 \cos^3 (\psi - t). \]

Let
\[ \bar{G}(\psi, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} r^2 \cos (\psi - t) \, dt. \]

This turns out to be \(-\frac{3}{8} r^2\), (independent of \( \psi \)). The averaged equation for \( \psi \) is
\[ \Psi' = -\frac{3}{8} R^2 \varepsilon \]
with solution \( \Psi = -\frac{3}{8} R^2 \varepsilon t \). Therefore the approximate angle is \( \Theta = -t - \frac{3}{8} R^2 \varepsilon t \) and we obtain the averaged solution
\[ \bar{x} = A \cos \left( -t - \frac{3}{8} A^2 \varepsilon t \right) = A \cos \left( t + \frac{3}{8} A^2 \varepsilon t \right). \]

This suggests that the period is approximately \( \frac{2\pi}{1 + \frac{3}{8} A^2 \varepsilon} \), though of course we have proved nothing rigorously yet.

We now state and prove our theorem concerning averaging. This is basically from Grimshaw.

**Theorem 1:** Suppose that there is a \( C_0 > 0, R_0, \) and \( \varepsilon_0 > 0 \) such that for \( 0 < \varepsilon < \varepsilon_0 \), the solution \( R(t) = R(t, \varepsilon) \) of (13) with \( R(0) = R_0 \) exists on the interval \([0, \frac{C_0}{\varepsilon}]\). Suppose that there is an \( M > 0 \) such that \( |R(t)| \leq M \) for all \( t \in [0, \frac{C_0}{\varepsilon}] \). (\( M \) is independent of \( \varepsilon \)). Then there is a \( C_1 \) (independent of \( \varepsilon \)) such that if \((r, \theta)\) solves (12) with \( r(0) = R(0) \), then \((r, \theta)\) exists on \([0, \frac{C_0}{\varepsilon}]\) and on this interval
\[ |R(t) - r(t)| \leq C_1 \varepsilon. \]
Thus, $R$ approximates $r$ on a time scale of $O\left(\frac{1}{\epsilon}\right)$.

(In the case of periodic solutions of (9), we expect the period to be close to $2\pi$, so an $O\left(\frac{1}{\epsilon}\right)$ time interval is more than enough to detect periodicity, as we will see rigorously below.)

Proof: This proof is fairly technical. One source of technicality is the assumed dependence of $f$ and $g$ on $\epsilon$, which leads to an assumption that $F$ and $G$ depends on $\epsilon$. Yet, in our examples, the $f$ and $g$ do not not depend on $\epsilon$. So I will give a proof in the case where $f, g$ and so $F, G$ are independent of $\epsilon$ and then outline what changes are needed for the general case. So we will assume that the equations in polar coordinates are

$$r' = \pm F(r, \theta)$$
$$\theta' = -1 + \pm G(r, \theta)$$

We then have

$$\tilde{F}(r) = \frac{1}{2\pi} \int_0^{2\pi} F(r, \theta) \, d\theta.$$  

Let $\hat{F}(r, \theta) = F(r, \theta) - \tilde{F}(r)$. Then $\hat{F}$ has period $2\pi$ in $\theta$, and

$$\int_0^{2\pi} \hat{F}(r, \theta) \, d\theta = 0.$$  

Let

$$H(r, \theta) = \int_0^\theta \hat{F}(r, s) \, ds.$$  

Then $H(r, 0) = H(r, 2\pi) = 0$ and $H$ has period $2\pi$ in $\theta$. Also, $H$ is bounded in the region $|r| \leq 2M$. (We need to choose a larger region than the bound assumed on $R$, since $r$ could be a little bigger than $R$. The bound $|r| \leq 2M$ is much larger than we actually need, but simple to work with.

Now let

$$\hat{r}(t) = r(t) + \pm H(r(t), \theta(t)).$$
Then

\[ r'(t) = r'(t) + \varepsilon \frac{\partial H}{\partial r} r'(t) + \varepsilon \frac{\partial H}{\partial \theta} \theta'(t) \]

\[ = \varepsilon F(r, \theta) + \varepsilon \frac{\partial H}{\partial r} \varepsilon F(r, \theta) + \varepsilon \hat{F}(r, \theta) (-1 + \varepsilon G(r, \theta)) \]

\[ = \varepsilon (\tilde{F} + \hat{F}) + \varepsilon^2 \frac{\partial H}{\partial r} F - \varepsilon \hat{F} + \varepsilon^2 \hat{F} G \]

\[ = \varepsilon F(r) + \varepsilon^2 Q(r, \theta), \]

where \( Q \) is continuous, bounded in \(|r| \leq 2M\), and has period \( 2\pi \) in \( \theta \). Also,

\[ \tilde{F}(r) = \hat{F}(\hat{r}) + \frac{\partial \hat{F}}{\partial r} (r^*) (r - \hat{r}) \]

for some \( r^* \) between \( r \) and \( \hat{r} \). Recall that \( \hat{r} = r + \varepsilon H \). Therefore

\[ \hat{r}' = \varepsilon \hat{F}(\hat{r}) + \varepsilon^2 P(r, \theta) \]

where \( P \) is \( 2\pi \)-periodic in \( \theta \) and bounded in the region \(|r| \leq 2M\). We compare this with the averaged equation

\[ R' = \varepsilon \tilde{F}(R), \]

and since the initial conditions of \( \hat{r} \) and \( R \) are close, we can expect that \( \hat{r} \) and \( R \) are close. This is what we show now.

Let \( z = \hat{r} - R \). Then

\[ z' = \varepsilon (\tilde{F}(\hat{r}) - \tilde{F}(R)) + \varepsilon^2 P \]

\[ z(0) = \varepsilon H(r(0), \theta(0)). \]

Now consider the Taylor series for \( \tilde{F} \). (We assume \( f \) is smooth enough for this to exist out to at least second order terms.) We have

\[ \tilde{F}(\hat{r}) = \tilde{F}(R) + \frac{\partial \tilde{F}}{\partial R}(R)(\hat{r} - R) + O(|\hat{r} - R|^2). \]

More precisely, there exist functions \( W(\hat{r}, R, \theta) \) and \( P_1(\hat{r}, R, \theta) \) such that

\[ z' = \varepsilon \frac{\partial \tilde{F}}{\partial R}(R) z + \varepsilon W(\hat{r}, R, \theta) + \varepsilon^2 P_1(\hat{r}, R, \theta) \]
and further, there is a $K > 0$ such that if $|\hat{r} - R|$ and $\varepsilon$ are sufficiently small, then

$$
|W(\hat{r}, R, \theta)| \leq K|\hat{r} - R|^2
$$

$$
|P_1(\hat{r}, R, \theta)| \leq K.
$$

We also have

$$
z(0) = \varepsilon H(r(0), \theta(0)).
$$

We choose $K$ as well so that $|H(r, \theta)| \leq K$ for $|r| \leq 2M$. Now let $\alpha = m\varepsilon KC_0e^{KC_0}$ where $C_0$ was given in the statement of the Theorem and

$$
m = \frac{8C_0 + 1}{5C_0}.
$$

(We shall need that later.) Then $|z(0)| = \varepsilon |H(r(0), \theta(0))| < \varepsilon K$. Since $mC_0e^{KC_0} > 1$ we see that $|z(0)| < \alpha$. Hence there is some $t_1 = t_1(\varepsilon)$ such that $|z| \leq \alpha$ on $[0, t_1]$. We must also assume that $\varepsilon t_1 \leq C_0$. Our goal is to show that we can choose $t_1 = \frac{C_0}{\varepsilon}$.

We then integrate the absolute value of the right side of (15), choosing $K$ in addition larger than $\max_{|r| \leq 2M} \left\{ \left| \frac{\partial \hat{r}}{\partial R} \right| \right\}$, and get

$$
|z(t)| \leq |z(0)| + K\varepsilon \int_0^t |z(s)| \, ds + K\alpha^2\varepsilon t + K\varepsilon^2 t.
$$

$$
\leq K\varepsilon \int_0^t |z(s)| \, ds + K\alpha^2\varepsilon t + K\varepsilon^2 t + K\varepsilon.
$$

on the interval $[0, t_1]$ where $|z| \leq \alpha$. This is valid because $|z|^2 = |\hat{r} - R|^2 \leq \alpha^2$ on $[0, t_1]$, and also $|r| \leq |\hat{r}| + K\varepsilon \leq R + |z| + K\varepsilon \leq M + |\alpha| + K\varepsilon < 2M$, if $\varepsilon$ is small enough.

On $[0, t_1]$ we have $\varepsilon t \leq C_0$, and therefore $K\varepsilon^2 t \leq K\varepsilon C_0$ and $K\alpha^2\varepsilon t \leq K\alpha^2 C_0$. Then from Gronwall’s lemma we get

$$
|z(t)| \leq (K\alpha^2 C_0 + \varepsilon KC_0 + K\varepsilon) e^{KC_0} \leq (K\alpha^2 C_0 + \varepsilon KC_0 + K\varepsilon) e^{KC_0}.
$$

(16)
Recall that we chose $\alpha$ to depend on $\varepsilon$. From the definitions of $\alpha$ and $m$ we see that the right side of (16) is equal to $(\alpha K_0 e^{K_0 t}) \alpha + \frac{5}{8} \alpha$. Because $\alpha = m \varepsilon C_0 K_0 e^{K_0 t} \to 0$ as $\varepsilon \to 0$, we see that for sufficiently small $\varepsilon$, the coefficient in parenthesis in the first term is less than $\frac{1}{8} \alpha$ and so the sum of the two terms is less than $\frac{3}{4} \alpha$.

Now recall the restrictions on $t_1$. They were that $|z| \leq \alpha$ on $[0, t_1]$ and that $t_1 \leq \frac{C_0}{\varepsilon}$. So we have proved that:

**If** $|z| \leq \alpha$ on $[0, t_1]$, and **if** $t_1 \leq \frac{C_0}{\varepsilon}$, **then** $|z| \leq \frac{3}{4} \alpha$ on $[0, t_1]$.

This implies that we can choose $t_1 = \frac{C_0}{\varepsilon}$.

Also, $|r - R| \leq |\hat{r} - R| + \varepsilon |H| \leq C_1 \varepsilon$ for some $C_1$ independent of $\varepsilon$.

I only discussed the case when $f$ does not depend on $\varepsilon$, and so $F$ and $G$ do not depend on $\varepsilon$. But the case where these functions do depend on $\varepsilon$ is not much harder. The only changes needed are around equation (14). That equation must become

$$\bar{F}(\hat{r}, \varepsilon) = \bar{F}(R, 0) + \frac{\partial \bar{F}}{\partial R}(R, 0)(\hat{r} - R) + \varepsilon \frac{\partial \bar{F}}{\partial \varepsilon}(R, 0) + O(\varepsilon^2 + |\hat{r} - R|^2)$$

and then the function $P_1$ in (15) will not be the same as the function $P$. But it will still be bounded and periodic in $|r| \leq 2M$.

This completes the proof.

We ask whether, for the example of Duffing’s equation this is enough to give a rigorous approximation to the period of the solution as a function of amplitude. We gave a nonrigorous argument that the period is approximately

$$T \simeq \frac{2\pi}{1 + \frac{3}{8} A^2 \varepsilon}.$$  

Using the Taylor series for $\frac{1}{1+x}$ this suggests that

$$T = 2\pi \left(1 - \frac{3}{8} A^2 \varepsilon + O(\varepsilon^2)\right)$$

where the $O(\varepsilon^2)$ term is positive, and therefore that a rigorous lower bound for $T$ might be

$$T > 2\pi \left(1 - \frac{3}{8} A^2 \varepsilon\right).$$
By definition of the polar coordinates, a period is defined by finding \( T \) such that \( \theta (T) = -2\pi \). (We know that \( \theta' \) is approximately \(-1\); hence the minus sign.)

For Duffing’s equation we found that the averaged equation was

\[
R' = 0.
\]

So solutions \( R = A \) exist on \((-\infty, \infty)\) and are bounded. Hence the hypotheses of Theorem 1 are satisfied for any \( C_0 \). Choosing \( C_0 = 1 \) we can assert that there is a \( C_1 \) such that \( |r - A| < C_1 \varepsilon \) on \([0, \frac{1}{2}]\). Turning to the \( \theta \) equation we then have

\[
\theta' = -1 - \varepsilon r^2 \cos^4 \theta.
\]

It is less confusing to set \( \psi (t) = -\theta (t) \) and get an estimate of \( T \) such that \( \psi (T) = 2\pi \). We have

\[
\psi' = 1 + \varepsilon^2 r^2 \cos \psi.
\]

From the bounds on \( r \) we get upper and lower bounds on \( \psi' \). An upper bound for \( \psi' \) should give us a lower bound for \( T \). An upper bound for \( \psi' \) is

\[
\psi' < 1 + \varepsilon B^2 \cos^4 \psi.
\]

where \( B = A + C_1 \varepsilon \).

**Homework:** Use the method of separation of variables to prove that if \( \lambda > \frac{3}{8} \), then for sufficiently small \( \varepsilon \),

\[
T > 2\pi (1 - \lambda A^2 \varepsilon).
\]

You may want to make use one of the inequalities

\[
1 - x + x^2 > \frac{1}{1 + x} > 1 - x.
\]

for small \( x > 0 \).

Using the same methods more precisely would give the result that

\[
\lim_{\varepsilon \to 0} \frac{T - 2\pi}{\varepsilon} = 2\pi \left( \frac{3}{8} A^2 \right).
\]

**Theorem 2** Suppose in addition to the hypotheses of Theorem 1 that \( R_0 \) is an isolated equilibrium point of the averaged equation (13). Suppose also that \( \frac{\partial F}{\partial R} |_{R_0} \neq 0 \).
Then for each sufficiently small \( \varepsilon > 0 \) there is a periodic solution of (12), such that

\[
\lim_{\varepsilon \to 0} r(t) = R_0(t) \text{ uniformly in } t.
\]

If \( \varepsilon \frac{\partial F}{\partial R}|_{R_0} < 0 \) then this solution is stable, while if \( \varepsilon \frac{\partial F}{\partial R}|_{R_0} > 0 \) it is unstable.

Proof:

Once again I will only discuss the case when \( f, \) and so \( F, G, \) are independent of \( \varepsilon \). For small \( \varepsilon, \theta' < -\frac{1}{2} \) on \([0, \frac{1}{\varepsilon}]\) and so on this interval we can think of \( r \) as a function of \( \theta \). We have

\[
\frac{dr}{d\theta} = \frac{\varepsilon F(r, \theta)}{-1 + \varepsilon G(r, \theta)} = \varepsilon K(r, \theta, \varepsilon).
\]

(18)

This system is non-autonomous, but the original system (12) is autonomous. Further, since \( \theta \) must cross some negative multiple of \( 2\pi \) in any time interval of length \( o \left( \frac{1}{\varepsilon} \right) \), and since \( F \) and \( G \) have period \( 2\pi \) in \( \theta \), we can assume that \( \theta(0) = 0 \). Assume that \( r(0) = r_0 \). We want to find \( r_0 \) such that \( r(-2\pi, r_0, \varepsilon) = r_0 \). But since the differential equation for \( r \) is \( 2\pi \)-periodic in \( \theta \), this is the same as saying \( r(2\pi, r_0, \varepsilon) = r_0 \). We denote the solution to (18) by \( r(\theta, r_0, \varepsilon) \). (This is an “abuse of notation”, since before we would have discussed “\( r(t, r_0, \theta_0, \varepsilon) \)”.)

For \( \varepsilon > 0 \) let

\[
N(r_0, \varepsilon) = \frac{1}{\varepsilon} (r(2\pi, r_0, \varepsilon) - r_0)
\]

We wish to find a solution to the equation

\[
N(r_0, \varepsilon) = 0,
\]

We first have to extend the definition of \( N \) to \( \varepsilon = 0 \) so that \( N \) is smooth and we can apply the implicit function theorem.

From (18) we have

\[
r(\theta, r_0, \varepsilon) = r_0 + \varepsilon \int_0^\theta K(r(s, r_0, \varepsilon), s, \varepsilon) \, ds
\]
and so

\[ N (r_0, \varepsilon) = \int_0^{2\pi} K (r (s, r_0, \varepsilon), s, \varepsilon) \, ds \]

\[ = \int_0^{2\pi} \frac{F(r(s, r_0, \varepsilon), s)}{-1 + \varepsilon G(r(s, r_0, \varepsilon), s)} \, ds. \]

Then it is obvious that \( N \) can be extended to \( \varepsilon = 0 \) with

\[ N (r_0, 0) = \int_0^{2\pi} -F(r_0, s) \, ds \]

\[ = -2\pi \bar{F}(r_0), \tag{19} \]

Hence

\[ N (R_0, 0) = 0. \]

and

\[ \frac{\partial N (R, 0)}{\partial r} \bigg|_{r=R_0} = -2\pi \frac{\partial \bar{F}}{\partial R} (R_0) \neq 0, \]

and therefore the implicit function theorem says that for small \( \varepsilon \) there is an \( r_0 (\varepsilon) \) such that \( N (r_0 (\varepsilon), \varepsilon) = 0 \). Also, \( r_0 (\varepsilon) \to R_0 \) as \( \varepsilon \to 0 \). This gives the desired family of periodic solutions and proves the existence part of the theorem.

For the stability, we consider an initial condition \( r_1 \neq r_0 (\varepsilon) \). We have

\[ r (\theta, r_1, \varepsilon) = r_1 + \int_0^\theta K (r (s, r_1, \varepsilon), s, \varepsilon) \, ds, \]

\[ r (-2\pi, r_1, \varepsilon) = r_1 + \int_{-2\pi}^0 K (r (s, r_1, \varepsilon), s, \varepsilon) \, ds \]

\[ = -\int_{-2\pi}^0 K (r (s, r_1, \varepsilon), s, \varepsilon) \, ds. \]

We let

\[ M (r_1, \varepsilon) = r (-2\pi, r_1, \varepsilon) - r_1 = -\int_{-2\pi}^0 K (r (s, r_1, \varepsilon), s, \varepsilon) \, ds. \]

\[ r (-2\pi, r_1, \varepsilon) = r_1 + \varepsilon M (r_1, \varepsilon). \]
Also,

\[ M (r_1, \varepsilon) = M (r_0 (\varepsilon), \varepsilon) + \frac{\partial M}{\partial r} (r^*, \varepsilon) (r_1 - r_0 (\varepsilon)) \]

\[ = \frac{\partial M}{\partial r} (r^*, \varepsilon) (r_1 - r_0 (\varepsilon)) \]

for some \( r^* \) between \( r_1 \) and \( r_0 (\varepsilon) \). We have

\[ M (R_0, 0) = - \int_{-2\pi}^{0} K (R_0, s, 0) ds = - \int_{0}^{2\pi} K (R_0, s, 0) ds = - N (R_0, 0) . \]

The second step is because \( K \) has period \( 2\pi \) in \( \theta \). If \( \frac{\partial F}{\partial r} (R_0, 0) = - \mu < 0 \), then from (19),

\[ \frac{\partial M}{\partial r} (R_0, 0) = - \frac{\partial N}{\partial r} (R_0, 0) = - \left( -2\pi \frac{\partial F}{\partial R} (R_0) \right) = -2\pi \mu < 0 \]

, and for sufficiently small \( |r_1 - R_0| + |\varepsilon| \) we have \( \frac{\partial M}{\partial r} (r^*, \varepsilon) \leq -2\pi \frac{\mu}{2} \). Hence for the case \( \varepsilon > 0 \) we have

\[ |r (-2\pi, r_1, \varepsilon) - r_0 (\varepsilon)| = \left| \left( 1 + \varepsilon \frac{\partial M}{\partial r} (r^*, \varepsilon) \right) (r_1 - r_0 (\varepsilon)) \right| \leq (1 - \pi \mu \varepsilon) |(r_1 - r_0 (\varepsilon))|. \]

Iterating this map of \( r_1 \rightarrow r (-2\pi, r_1, \varepsilon) \) over and over, we see that the solution tends to \( r_0 (\varepsilon) \), at an exponential rate. More specifically, we have \( |r (-2k\pi) - r_0| \leq (1 - \pi \mu \varepsilon)^k |r (0) - r_0| \), where \( r \) is a function of \( \theta \). Since \( \theta' = -1 + O (\varepsilon) \), we can, approximately, replace \( \theta \) with \(-t\). We see that with \( r \) a function of \( t \), \( |r (2k\pi) - r_0| \leq (1 - \pi \mu \varepsilon)^k |r (0) - r_0| \) . Taking logs,

\[ \log |r (2k\pi) - r_0 (\varepsilon)| \leq \log |r (0) - r_0| + k \log (1 - \pi \mu \varepsilon) \sim -k \pi \mu \varepsilon \]

as \( k \rightarrow \infty \). Hence \( |r (t) - r_0| = O (e^{\frac{-\mu \pi}{2} t}) \).

Remark: That word “approximately” above means that this is not rigorous! As an exercise you should try to fix this point.

Applications:

It is easy to apply this to the van der Pol equation. We have, from (12),

\[ f (x, y) = (1 - x^2) y \]
and so

$$F(r, \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} (1 - r^2 \cos^2 \theta) r \sin^2 \theta d\theta$$

$$= \frac{1}{2\pi} \left( \pi r - \frac{1}{4} r^3 \pi \right).$$

We see that the desired zero is $R_0 = 2$, confirming that there is a family of periodic solutions for $\varepsilon > 0$ which tends to the circle of radius 2 in the phase plane as $\varepsilon \to 0$. This is finally a rigorous proof of this result.

We also get information about other solutions, besides the periodic one. The averaged equation is

$$R' = \varepsilon \left( \frac{R}{2} - \frac{1}{8} R^3 \right)$$

with solution $R = \frac{2}{\sqrt{1+4\varepsilon}}$ and so every solution of the average equation satisfies

$$\lim_{t \to \infty} R(t) = 2.$$

This is consistent with stability of the periodic solution, but we have not proved this, since our theorem only says that the averaged solution is an approximation of the periodic solution for times of order $O \left( \frac{1}{\varepsilon} \right)$. However a stronger statement of the theorem, given in GH, pg. 168, says that solutions on the stable manifold of the periodic solution are approximated by solutions on the stable manifold of the fixed point of the averaged equation uniformly on $[0, \infty)$. The complete proof is not given there however, and reference is made to an article from 1982, which apparently was the first complete proof.

Averaging gets more interesting when we turn to non-autonomous equations, namely equations of forced oscillations. The general form to be considered now is

$$x'' + x = \varepsilon g(x, x', t, \varepsilon)$$

(20)

where $g$ is periodic in $t$ with period $T$.

The difficulty of finding approximate solutions of (20) depends on how near to “resonance” it is. We have not defined resonance for a nonlinear system, but by considering the linear case we can see that solutions are more difficult if $T = 2\pi$, since that is the natural frequency of the unforced system.
Another point to be kept in mind is that $g$ may contain a linear part, say $\mu x$ which will affect the period. It is convenient to write the equation as

$$x'' + (1 + \kappa \varepsilon)^2 x = \varepsilon h(x, x', t, \varepsilon).$$

Consider, for example, the linear nonautonomous case where

$$h(x, x', t, \varepsilon) = \cos t.$$  

Thus we have

$$x'' + (1 + \kappa \varepsilon)^2 x = \varepsilon \cos t.$$  

We see that if $\kappa \neq 0$, then for any fixed $\varepsilon$, there is no resonance, since the forcing frequency is different from the natural frequency. But it is not clear what happens as $\varepsilon \to 0$, since in this limit the natural frequency is becoming closer to the forcing frequency, but on the other hand the forcing amplitude is going to zero.

We can use the variation of parameters formula to show that this equation has a unique solution with the period, $2\pi$, of the forcing function. This solution is most easily found by trying a solution $A \cos t + B \sin t$. We find that

$$x = \frac{\varepsilon}{(1 + \kappa \varepsilon)^2 - 1} \cos t.$$  

Expanding the coefficients in terms of $\varepsilon$, we find that as $\varepsilon \to 0$ we approach the solution

$$\frac{1}{2 \kappa} \cos t.$$  

It is common to study the amplitude of the solution as a function of the “tuning parameter” $\kappa$. This is the curve $y = \left| \frac{1}{2 \kappa} \right|$.

Now we want to see how this curve changes if we add two terms: a small damping term, and a small nonlinear term.

We can start by adding a damping term to the linear equation, giving

$$x'' + \varepsilon \delta x' + (1 + \kappa \varepsilon)^2 x = \varepsilon \cos t.$$  

As $\varepsilon \to 0$ the unique periodic solution now approaches

$$\frac{2 \kappa}{4 \kappa^2 + \delta^2} \cos t + \frac{\delta}{4 \kappa^2 + \delta^2} \sin t.$$
For positive $\delta$ the amplitude is now bounded as $\kappa \to 0$, approaching $\frac{1}{\delta}$.

Next, we set the damping equal to zero and add a cubic term. We also follow Grimshaw and put the term keeping the linearized equation away from resonance on the right hand side. This $\kappa$ is not quite the same as the $\kappa$ above, which would have included $\kappa^2$ if we had put the $\kappa$ terms on the right, but it still is “near resonance” as $\varepsilon \to 0$. We wish to use averaging to find the limiting amplitude response curve. We will now allow for a variable strength forcing, though still proportional to $\varepsilon$.

$$\ddot{x} + x = \varepsilon \left(G_0 \cos t - \kappa x + x^3\right),$$

(21)

We will see the effect of the nonlinear term $\varepsilon x^3$.

We first try the simple expansion

$$x = x_0 + \varepsilon x_1 + ...$$

and get

$$\ddot{x}_0 + x_0 = 0$$

$$\ddot{x}_1 + x_1 = (G_0 - A\kappa) \cos t + A^3 \cos^3 t$$

Going through our usual procedure to eliminate resonance gives

$$G_0 - \kappa A + \frac{3}{4} A^3 = 0.$$  

This gives at least one, and perhaps three values of $A$, depending on $\kappa$. Notice that we don’t have to use multiple scales here, assuming that $G_0$ and $\kappa$ are non-zero. But this is only at the first stage. The naive method might fail when we look for $x_2$. (I haven’t tried it!)

Now let’s try averaging on this problem. It is of a different form from before, and the averaging must now be in $t$ instead of the phase variable $\theta$. This requires a different approach.

The general averaging theorem for non-autonomous perturbations is given in GH, page 168. For the special case of (21) we proceed as follows.
We write this as a system

\[ \begin{align*}
x' &= y \\
y' &= -x + \varepsilon g(x, y, t, \varepsilon)
\end{align*} \tag{22}\]

and look for a solution of the form

\[ \begin{align*}
x &= a(t) \cos t + b(t) \sin t \\
y &= -a(t) \sin t + b(t) \cos t.
\end{align*} \tag{23}\]

The hope is that because \(x^2 + y^2 = 2\varepsilon y g(x, y, t, \varepsilon)\), and so the amplitude of the oscillation changes slowly, we will find that \(a(t)\) and \(b(t)\) change slowly. The solution is then an oscillation with period equal to \(2\pi\) and with slowly varying amplitude. The period must be exactly \(2\pi\) because \(g\) has period \(2\pi\).

To find equations satisfied by \(a\) and \(b\) we differentiate (23) and use (22). Thus,

\[ \begin{align*}
x' &= a' \cos t - a \sin t + b' \sin t + b \cos t = y = -a \sin t + b \cos t \\
y' &= -a' \sin t - a \cos t + b' \cos t - b \sin t = -x + \varepsilon g = -a \cos t - b \sin t + \varepsilon g(x, y, t, \varepsilon)
\end{align*} \]

where we substitute the expressions on the right of (23) for \(x\) and \(y\) in this last equation. Solving for \(a'\) and \(b'\) gives

\[ \begin{align*}
a' &= -\varepsilon \sin t \ g(x, y, t, \varepsilon) = \varepsilon F(a, b, t, \varepsilon) \\
b' &= \varepsilon \cos t \ g(x, y, t, \varepsilon) = \varepsilon G(a, b, t, \varepsilon).
\end{align*} \tag{24}\]

Thus, as expected, \(a'\) and \(b'\) are \(O(\varepsilon)\). Note that \(F\) and \(G\) are of period \(2\pi\) in \(t\).

We now use averaging by letting

\[ \begin{align*}
F(a, b, \varepsilon) &= \frac{1}{2\pi} \int_0^{2\pi} F(a, b, t, \varepsilon) \, dt \\
G(a, b, \varepsilon) &= \frac{1}{2\pi} \int_0^{2\pi} G(a, b, t, \varepsilon) \, dt.
\end{align*} \]

The averaged equations are then

\[ \begin{align*}
A' &= \varepsilon \bar{F}(A, B, 0) \\
B' &= \varepsilon \bar{G}(A, B, 0). \tag{25}
\end{align*} \]
Before proceeding, note the hierarchy of averaged equations. For a single non-autonomous equation,

\[ x' = \varepsilon f(x, t, \varepsilon) \]

the averaged equation is a first order autonomous equation

\[ y' = \varepsilon F(y, 0). \]

For a second order autonomous equation

\[ x'' + x = \varepsilon f(x, x', \varepsilon) \]

the averaged equation was a first order autonomous equation (for the amplitude)

\[ R' = \varepsilon F(R, 0). \]

For a second order non-autonomous equation

\[ x'' + x = \varepsilon f(x, x', t, \varepsilon) \]

(equivalent to a nonautonomous system of two first order equations) the average system is a system of two first order autonomous equations, (25). The advantage of this over the original nonautonomous system is that it can be studied in the phase plane.

We have the following theorems:

**Theorem 3.** Suppose that there is a \( C_0 \) and a point \((A_0, B_0)\) such that for each sufficiently small \( \varepsilon \) \( > 0 \) the solution \((A, B)\) of (25) with \( A(0) = A_0, B(0) = B_0 \) exists on the interval \([0, \frac{C_0}{\varepsilon}]\). Suppose that \(|A| + |B| \leq M\) on this interval, where \( M \) is a constant independent of \( \varepsilon \). Then there is a \( C_1 > 0 \) such that if \((a, b)\) solve (24) and \( a(0) = A(0), b(0) = B(0) \), then \((a, b)\) exists on \([0, \frac{C_0}{\varepsilon}]\) and \(|a - A| + |b - B| \leq C_1 \varepsilon\) on \([0, \frac{C_0}{\varepsilon}]\).

**Theorem 4.** Suppose in addition that \((A_0, B_0)\) is an isolated equilibrium point of the averaged equation (25). Suppose also that \( \det \begin{bmatrix} \frac{\partial F}{\partial A} & \frac{\partial F}{\partial B} \\ \frac{\partial G}{\partial A} & \frac{\partial G}{\partial B} \end{bmatrix} \bigg|_{(A_0, B_0)} \neq 0 \). Then for each sufficiently small \( \varepsilon \) there is a periodic solution \((a, b)\) of (24), such that \( \lim_{\varepsilon \to 0} (a, b) = (A_0, B_0) \) uniformly in \( t \). This periodic solution has the same stability as the equilibrium point does for (25).
The proofs are similar to those of Theorems 1 and 2. See Grimshaw, pg. 225.

We now consider the forced Duffing equation, that is

\[ g(x, y, t, \varepsilon) = G_0 \cos t - \kappa x + x^3. \]

Then,

\[
F(a, b, t, \varepsilon) = -\sin t \left( G_0 \cos t - \kappa (a \cos t + b \sin t) + (a \cos t + b \sin t)^3 \right)
\]

\[
G(a, b, t, \varepsilon) = \cos t \left( G_0 \cos t - \kappa (a \cos t + b \sin t) + (a \cos t + b \sin t)^3 \right)
\]

\[
\frac{1}{2\pi} \int_0^{2\pi} (-\sin t) \left( G_0 \cos t - \kappa (a \cos t + b \sin t) + (a \cos t + b \sin t)^3 \right) dt
\]

\[ = \frac{1}{2} \left( -\frac{3}{4} b^3 + b\kappa - \frac{3}{4} a^2 b \right) \]

\[
\frac{1}{2\pi} \int_0^{2\pi} (\cos t) \left( G_0 \cos t - \kappa (a \cos t + b \sin t) + (a \cos t + b \sin t)^3 \right) dt
\]

\[ = \frac{1}{2} \left( G_0 - \kappa a + \frac{3}{4} a^3 + \frac{3}{4} ab^2 \right) \]

Hence the averaged equations are

\[
A' = \frac{1}{2} \varepsilon B \left( \kappa - \frac{3}{4} (A^2 + B^2) \right) \quad \text{(26)}
\]

\[
B' = \frac{1}{2} \varepsilon \left( G_0 - A \left( \kappa - \frac{3}{4} (A^2 + B^2) \right) \right).
\]

We first look for equilibrium points. From the first equation, either \( B = 0 \) or \( \kappa = \frac{3}{4} (A^2 + B^2) \). But the second alternative gives no equilibrium with \( G_0 \neq 0 \), so we have \( B_0 = 0 \), and then \( A_0 \) satisfies

\[
G_0 = A_0 \kappa - \frac{3}{4} A_0^3. \quad \text{(27)}
\]
Once again we study how the amplitude depends on $\kappa$, the “response diagram” for amplitude. It is found in practically every book on nonlinear oscillations. See Grimshaw, page 204. It gives the amplitude of the periodic solution as a function of the tuning parameter $\kappa$. One feature to notice is that for large enough $\kappa$ there are two solutions $A$, and the maximum solution tends to infinity as $\kappa \to \infty$. This is for the equation (24), which has no term for damping of the oscillation.

Notice also that the amplitude is actually $|A|$. If $A < 0$ but $G_0 > 0$ then we solve (27) for small $|A|$ by taking $\kappa < 0$. Thus there are two branches to the curve. We can study the graph by solving for $\kappa$:

$$\kappa = \frac{G_0 + \frac{3}{4}A^3}{A}$$

for $A \neq 0$. The graph will plot $|A|$ on the positive vertical axis vs $\kappa$ on the horizontal axis. There are two branches. On one, $\kappa \to -\infty$ as $|A| \to 0$ and $\kappa \to \infty$ as $|A| \to \infty$. On the other, $\kappa \to \infty$ both as $|A| \to 0$ and as $|A| \to \infty$.

If a small damping term is added, giving

$$x'' + \varepsilon \delta x' + x = \varepsilon \left(G_0 \cos t - \kappa x + x^3\right)$$

then we put it on the right side, adding a term $\delta (-a \sin t + b \cos t)$ to the expression for $F(a, b, t, \varepsilon)$. The equation for steady states becomes $A^2 \left(\delta^2 + (\kappa - \frac{3}{4}A^2)^2\right) = G_0^2$. One difference is that $A$ is bounded, since $A^2 \leq G_0^2/\delta^2$. There is only one branch, continuous but not single valued in either $A$ or $\kappa$. Thus the damping keeps the amplitude bounded independent of $\kappa$.

We then want to analyze the stability of the periodic solutions, which our theorem says is the same as the stability of the equilibrium point, if it is hyperbolic (no eigenvalues with zero real part). For the $\delta = 0$ case we linearize (26) around a point where $G_0 = A\kappa - \frac{3}{4}A^3$, $B = 0$. We get the linearized matrix

$$
\begin{pmatrix}
0 & \frac{1}{2\varepsilon} (\kappa - \frac{3}{4}A^2) \\
\frac{1}{2\varepsilon} (-\kappa + \frac{9}{4}A^2) & 0
\end{pmatrix}
$$

For some values of $A$, and the corresponding $\kappa = \frac{G_0 + \frac{3}{4}A^3}{A}$, the roots will be pure imaginary and we don’t learn about stability from the linearization. However in the region where the two terms in the matrix have the same sign we will have one real positive eigenvalue and one real negative eigenvalue and so the equilibrium point
of the averaged equation, and the periodic solution of the unaveraged equation, are unstable. This is the region
\[ \frac{3}{4} A^2 < \kappa < \frac{9}{4} A^2 \]
between two parabolas. It can be checked that the parabola on the right intersects the right branch of equilibrium points in the \((\kappa, |A|)\) plane exactly where this branch turns around.

For \(\delta > 0\) it turns out that we can show that the open region outside of the instability region is stable.

The big advantage of averaging is that it gives information about more than just periodic solutions. This is seen in Theorems 1 and 3. The information is only over an interval of length \(O \left( \frac{1}{2} \right)\), but that is still useful. In fact, the averaging theorem in GH, pg. 168, gives further results that hold on \([0, \infty)\), for special solutions, those on the “stable manifold” of a periodic solution. This is a concept we have yet to discuss, but we will at a later point.

To get information about other solutions besides periodic ones we consider the averaged system (26). Before we looked for equilibria, which correspond to periodic solutions. Now we look at other orbits in the \((A, B)\) phase plane. These will approximate solutions of (24) on a time scale \(O \left( \frac{1}{2} \right)\).

In fact, the \((A, B)\) phase plane is on the cover of Grimshaw’s book, as well as on page 230. It appears in GH on page 183. It will be noted that there are two periodic orbits, and three families of periodic solutions.

To analyze the system (26) it is useful to once again use polar coordinates. We set
\[ A = R \cos \Phi \]
\[ B = R \sin \Phi. \]

Then we find that
\[ R' = \frac{1}{2} \varepsilon G_0 \sin \Phi \]
(28)
\[ \Phi' = \frac{1}{2R} \varepsilon \left[ G_0 \cos \Phi \right] - R \left( \kappa - \frac{3}{4} R^2 \right). \]

28
It is difficult to deal with either (25) or (28) unless, by some chance, there is a conserved quantity, say \( Q(R, \Phi) \) in the variables of (28) which is constant along solutions. We can have some hope for this because the original system,

\[
x'' + x = \varepsilon \left( G_0 \cos t - \kappa x + x^3 \right)
\]

has a conserved quantity if \( G_0 = 0 \), namely

\[
E = \frac{1}{2} x'^2 + \frac{1}{2} x^2 (1 + \kappa) - \frac{1}{4} x^4.
\]

Since in some sense we averaged the term \( G_0 \cos t \), and it has zero average, we may hope that there is still a conserved quantity.

It is a bit of work to convert the function \( E \) into the new variables, and as far as I took it, I didn’t see that it led to anything. But somehow it was discovered that

\[
\frac{1}{2} \kappa R^2 - \frac{3}{16} R^4 - G_0 R \cos \Phi
\]

is constant. This leads to the fact that all the orbits, in either the \((R, \Phi)\) or \((A, B)\) planes, are closed curves or else homoclinic orbits or pairs of heteroclinic orbits as in the pendulum phase plane.

The number of equilibrium points is seen from (27) to be between 1 and 3, depending on \( \kappa \) and \( G_0 \). Choosing values where there are three, it is found that two are centers and one a saddle, and that there are two homoclinic orbits tending to the saddle. One gets a good picture using xpp on the computer.

Theorem 4 tells us that a fixed point of (25) corresponds to a family of periodic solutions to (24) for small \( \varepsilon \), if the appropriate Jacobian determinant is nonzero. It is important to note, however, that a periodic solution of (25) does not necessarily correspond to a periodic solution of (24). This is for two reasons: First, solutions of (25) only approximate solutions of (24). Second, the period of a solution in the \((A, B)\) phase plane will probably not be \( 2\pi \). There is no reason for it to have this value, since the \((A, B)\) phase plane is autonomous, and the periodic solutions have many different periods. The superimposed oscillation of period \( 2\pi \) due to the \( 2\pi \) periodicity of \( F \) and \( G \), is likely to be incommensurate with the period in the \(A, B\) plane, so that when the solution in the \((A, B)\) plane returns to its starting point,
the values of \( \sin t \) and \( \cos t \), in the formula \( x = a \cos t + b \sin t \), where \(|a - A|\) and \(|b - B|\) are small, will probably not be back to their starting points.

Relation to Poincaré maps.

There are two notions of Poincaré map which come up frequently in oscillation theory. The first applies to autonomous systems. While the terminology “Poincaré map” may not have been used, the concept was used in the proof of the Poincaré-Bendixson theorem. We considered a two dimensional system,

\[
x' = f(x)
\]

(29)

where all vectors were two dimensional, and assumed there was a periodic solution \( p(t) \), say of period \( T \). We then considered a transversal, \( M \), which was a line segment intersecting the orbit of \( p \) and not tangent to this orbit at the point of intersection. We then considered (in the main theorem about asymptotic orbital stability of \( p \)), the “first return” map, following the solution through an initial point \( \eta \) until the next point \( P(\eta) \) where this solution again intersected \( M \). This map, \( \eta \rightarrow P(\eta) \) is called a Poincaré map for the system (29).

There is a second concept of Poincaré map, which is useful in periodic systems of the form

\[
x' = f(x,t)
\]

(30)

where \( f(x,t+T) = f(x,t) \) for all \((x,t)\). This is a map defined on \( \mathbb{R}^n \) where \( n \) is the dimension of \( x \). (By contrast, the Poincaré map for (29) is a one dimensional map if (29) is a two dimensional autonomous system.)

The Poincaré map for (30) is the map

\[
x_0 \rightarrow x(T,x_0)
\]

where the right side denotes the value at time \( T \) of the solution with initial value \( x(0) = x_0 \).

This can be interpreted as being of the same dimension as the previous kind of map if we write the non-autonomous system (30) on \( n \) equations as an autonomous system of \( n + 1 \) equations, namely

\[
\begin{align*}
y' &= f(y,s) \\
s' &= 1
\end{align*}
\]

(31)
with \( s(0) = 0 \). In this case \( s(t) = t \) and the systems (30) and (31) are equivalent. Then the Poincaré map consists of a map from the \( n \)-dimensional subspace \( s = 0 \) of the \( (n + 1) \)-dimensional \((x,s)\) space phase space of (31) to the \( n \)-dimensional subspace \( s = T \). Thus, like the map for (29), the Poincaré map is on a space of one less dimension than the phase space for the autonomous system. However in this case, there is no necessary relation to a periodic solution.

If, however, (30) does have a periodic solution \( p(t) \), then \( p(0) \) will be a fixed point for the Poincaré map for (31).

It is useful to look at the following example.:

\[
x' = \varepsilon (-x + \cos^2 t)
\]

and

\[
y' = \varepsilon (-y + \frac{1}{2})
\]

The general solutions are

\[
x(t) = \frac{\varepsilon^2 \cos 2t + 2\varepsilon \sin 2t}{2\varepsilon^2 + 8} + \frac{1}{2} + ce^{-et}
\]

for (32), and

\[
y(t) = \frac{1}{2} + ce^{-et}
\]

for (33). The first equation, (32) has a unique periodic solution \( p(t, \varepsilon) \), obtained by setting \( c = 0 \) in the formula for \( x(t) \). It is a non-autonomous equation, so we consider the Poincaré map, which is the “time 2\( \pi \)” map, namely \( x(0) \to x(2\pi) \). We have \( p(0, \varepsilon) = \frac{\varepsilon^2}{\varepsilon^2 + 8} + \frac{1}{2} = p_0 \) and this is the fixed point of the Poincaré map. Nearby points will start at \( p_0 + c \) for some small \( c \), and the Poincaré map is

\[
x(0) = p_0 + c \to x(2\pi) = p_0 + ce^{-2\pi}.
\]

In other words,

\[
P_\varepsilon(x_0) = \frac{\varepsilon^2}{\varepsilon^2 + 8} + \frac{1}{2} + \left( x_0 - \left( \frac{\varepsilon^2}{\varepsilon^2 + 8} + \frac{1}{2} \right) \right) e^{-2\pi}.
\]
The second equation is autonomous. In this case we again consider the time $2\pi$ map. This time of $2\pi$ is not intrinsic to the equation. We choose it because of its relation to (32). The map is

$$P_0(y_0) = \frac{1}{2} + \left(y_0 - \frac{1}{2}\right) e^{-2\pi}.$$  

We therefore see that in this case

$$P_\varepsilon(z) - P_0(z) = O(\varepsilon^2).$$

Theorem 3 concerning the method of averaging in a forced oscillation asserts that solutions of the original equations and the averaged form which start at the same time stay close together over an $O\left(\frac{1}{\varepsilon}\right)$ time scale. This is far longer than one period, so these theorems can be interpreted as saying that the time $T$ Poincaré maps for the two systems are very close – within $O(\varepsilon)$. The estimate above is smaller because we are considering a fixed interval of length $T$, not an interval of length $O\left(\frac{1}{\varepsilon}\right)$.

Theorem 4 discusses the relation between a hyperbolic fixed point for the averaged system and a periodic solution of the original forced oscillation system. Recall the systems (24) and (25), which were

$$a' = \varepsilon F(a, b, t, \varepsilon)$$
$$b' = \varepsilon G(a, b, t, \varepsilon).$$

and

$$A' = \varepsilon \bar{F}(A, B, 0)$$
$$B' = \varepsilon \bar{G}(A, B, 0).$$

The first system has a natural time $T$ Poincaré map. The initial value of a periodic solution is a fixed point for this map. The second system has a time $T$ Poincaré map also, the time being chosen because it is appropriate for the first system. An equilibrium point for the second, autonomous system, is a fixed point for its Poincaré map. The averaging theorem implies that these maps are very close to each other. But before we can go into further detail about this, we need to have a diversion into the study of Poincaré maps, and indeed, $n$-dimensional maps in general.