12

Advanced Averaging

In this chapter, we will look at a number of useful and important extensions of averaging. Also, at the end there will again be a discussion of timescales.

12.1 Averaging over an Angle

In oscillations, we can identify amplitudes and angles and also other quantities such as energy, angular momentum, etc. An angle can often be used as a time-like quantity that can be used for averaging. A simple example is the Duffing equation in Example 11.5.

Example 12.1
Consider the Duffing equation with positive damping and basic (positive) frequency ω,

$$\ddot{x} + \varepsilon\mu\dot{x} + \omega^2 x + \varepsilon\gamma x^3 = 0.$$

Using amplitude-phase variables r, ψ, we found slowly varying equations that could be averaged. Instead, we put $\dot{x} = \omega y$ and propose an amplitude-angle transformation of the form

$$x(t) = r(t)\sin\phi(t),\ y(t) = r(t)\cos\phi(t). \tag{12.1}$$

When transforming $x, \dot{x}(y) \to r, \phi$, we use the differential equation and the relation between $\dot{x}$ and y to find

$$\dot{r} = \varepsilon\frac{\cos\phi}{\omega}(-\mu\omega r\cos\phi - \gamma r^3\sin^3\phi),$$
$$\dot{\phi} = \omega - \varepsilon\frac{\sin\phi}{\omega r}(-\mu\omega r\cos\phi - \gamma r^3\sin^3\phi).$$

These equations are 2π-periodic in ϕ, and the equation for r looks like the standard form for averaging except that time is not explicitly present. We average the first equation over ϕ to find

$$\dot{r}_a = -\frac{1}{2}\varepsilon\mu r_a$$

with solution $r_a(t) = r(0)\exp(-\frac{1}{2}\varepsilon t)$.

What does $r_a(t)$ represent? We have averaged over ϕ but are interested in the behaviour with time. Below we will formulate a theorem that states that $r_a(t)$ is an $O(\varepsilon)$ approximation of $r(t)$ on the timescale $1/\varepsilon$ if ω is bounded away from zero ($\omega = O_s(1)$).

The last condition is quite natural. In averaging over ϕ, we treated ϕ as a time-like variable. This would not hold in an asymptotic sense if for instance ω is $O(\varepsilon)$. In such a case, a different approach is needed.

We will now look at a classical example, a linear oscillator with slowly varying (prescribed) frequency.

Example 12.2
Consider the equation

$$\ddot{x} + \omega^2(\varepsilon t)x = 0.$$

We put $\dot{x} = \omega(\varepsilon t)y, \tau = \varepsilon t$, and use transformation (12.1) to find

$$\begin{aligned}
\dot{r} &= -\varepsilon\frac{1}{\omega(\tau)}\frac{d\omega}{d\tau}r\cos^2\phi,\\
\dot{\phi} &= \omega(\tau) + \varepsilon\frac{1}{\omega(\tau)}\frac{d\omega}{d\tau}\sin\phi\cos\phi,\\
\dot{\tau} &= \varepsilon.
\end{aligned}$$

If we assume that $0 < a < \omega(\tau) < b$ (with a, b constants independent of ε), we have a three-dimensional system periodic in ϕ with two equations slowly varying. The function $\omega(\tau)$ is smooth and its derivative is bounded.

Averaging the slowly varying equations over ϕ, we obtain

$$\begin{aligned}
\dot{r}_a &= -\varepsilon\frac{1}{2\omega(\tau)}\frac{d\omega}{d\tau}r_a,\\
\dot{\tau} &= \varepsilon.
\end{aligned}$$

We write τ instead of τ_a, as the equation for τ does not change. From this system, we get

$$\frac{dr_a}{d\tau} = -\frac{1}{2\omega(\tau)}\frac{d\omega}{d\tau}r_a,$$

which can be integrated to find

$$r_a(\tau) = \frac{r(0)\sqrt{\omega(0)}}{\sqrt{\omega(\tau)}}.$$

Note that the quantity $r_a(\varepsilon t)\sqrt{\omega(\varepsilon t)}$ is conserved in time. The same techniques can be applied to equations of the form

$$\ddot{x} + \varepsilon\mu\dot{x} + \omega^2(\varepsilon t)x + \varepsilon f(x) = 0.$$

The application is straightforward and we leave this as an exercise.

We now summarise the result behind these calculations:

Theorem 12.1
Consider the system

$$\begin{aligned}\dot{x} &= \varepsilon X(\phi, x) + \varepsilon^2 \cdots ,\\ \dot{\phi} &= \omega(x) + \varepsilon \cdots ,\end{aligned}$$

where the dots stand for higher-order terms. We assume that $x \in D \subset \mathbb{R}^n$, ϕ is one-dimensional, and $0 < \phi < 2\pi$. Averaging over the angle ϕ produces

$$X^0(y) = \int_0^{2\pi} X(\phi, y) d\phi.$$

Assuming that

- the right-hand sides of the equations for x and ϕ are smooth;
- the solution of
$$\dot{y} = \varepsilon X^0(y),\ y(0) = x(0)$$
is contained in an interior subset of D;
- $\omega(x)$ is bounded away from zero by a constant independent of ε,

then $x(t) - y(t) = O(\varepsilon)$ on the timescale $1/\varepsilon$.

What happens if the frequency in Example 12.2 is not bounded away from zero? We consider a well-known example that can be integrated exactly.

Example 12.3
A spring with a stiffness that wears out in time is described by

$$\ddot{x} + e^{-\varepsilon t}x = 0,\ \ x(0) = 0, \dot{x}(0) = 1.$$

In the analysis of Example 12.2, we considered such an equation with the condition that the variable frequency $\omega(\varepsilon t)$ is bounded away from zero. On the *timescale* $1/\varepsilon$, this is still the case, so we conclude from Example 12.2 that the amplitude $r(t)$ on this timescale is approximated by

$$r_a(\varepsilon t) = \frac{r(0)\sqrt{\omega(0)}}{\sqrt{\omega(\varepsilon t)}}.$$

With $r(0) = 1, \omega(\varepsilon t) = e^{-\frac{1}{2}\varepsilon t}$, we find

$$r(t) = e^{\frac{1}{4}\varepsilon t} + O(\varepsilon)$$

on the timescale $1/\varepsilon$. Of course, this result may not carry through on longer timescales.

The exact solution can be obtained by transforming $s = 2\varepsilon^{-1}e^{-\varepsilon t/2}$, which leads to the Bessel equation of index zero and the solution

$$x(t) = \frac{\pi}{\varepsilon}Y_0\left(\frac{2}{\varepsilon}\right)J_0\left(\frac{2}{\varepsilon}e^{-\frac{1}{2}\varepsilon t}\right) - \frac{\pi}{\varepsilon}J_0\left(\frac{2}{\varepsilon}\right)Y_0\left(\frac{2}{\varepsilon}e^{-\frac{1}{2}\varepsilon t}\right),$$

where J_0, Y_0 are the well-known Bessel functions of index zero. On the timescale $1/\varepsilon$, all of the arguments are large and we can use the known behaviour of Bessel functions for arguments tending to infinity. We find that the solution is approximated by

$$x(t) = e^{\frac{\varepsilon t}{4}} \sin\left(\frac{2}{\varepsilon}(1 - e^{-\varepsilon t/2})\right) + O(\varepsilon).$$

Naturally, this agrees with the results obtained by averaging. If time runs beyond $1/\varepsilon$, the arguments of the Bessel functions J_0 and Y_0 tend to zero. Using the corresponding known expansions, we find that the solution behaves as $c_1 t + c_2$ with

$$c_1 = \sqrt{\frac{\varepsilon}{\pi}}\cos\left(\frac{2}{\varepsilon} - \frac{\pi}{4}\right),\ c_2 = \frac{\pi}{\varepsilon}\left(\sin\left(\frac{2}{\varepsilon} - \frac{\pi}{4}\right) - \frac{2}{\pi}\cos\left(\frac{2}{\varepsilon} - \frac{\pi}{4}\right)(\ln\frac{1}{\varepsilon} + \gamma)\right),$$

where γ is Euler's constant. Interestingly, on timescales larger than $1/\varepsilon$, there are no oscillations anymore, while the velocity c_2 becomes large with ε.

In the examples until now, we have introduced a different way of averaging but not a real improvement on elementary averaging as discussed in the preceding chapter. An improvement can arise when the equation for ϕ is rather intractable and we are satisfied with an approximation of the quantity x.

As we have seen, interesting problems arise when $\omega(x)$ is not bounded away from zero. In the next section, this will become a relevant issue in problems with more angles. To prepare for this, we consider an example from Arnold (1965).

Example 12.4

Consider the two scalar equations

$$\dot{x} = \varepsilon(1 - 2\cos\phi),$$
$$\dot{\phi} = x.$$

If $x(0) > 0$, independent of ε, $\omega(x) = x$ will remain bounded away from zero and averaging produces

$$\dot{y} = \varepsilon,\ y(0) = x(0),$$

so that $y(t) = x(0) + \varepsilon t$ and $x(t) - y(t) = O(\varepsilon)$ on the timescale $1/\varepsilon$. What happens if $x(0) \leq 0$? The system is conservative (compute the divergence or differentiate the equation for ϕ to eliminate x), and the original system has two stationary solutions (critical points of the right-hand side vector field): $(x, \phi) = (0, \frac{\pi}{3})$ and $(x, \phi) = (0, \frac{5\pi}{3})$. So for these solutions the expression $y(t) = x(0) + \varepsilon t$ is clearly not an approximation. However, this expression represents an approximation of $x(t)$ if $x(0) < 0$ until $x(t)$ enters a neighbourhood of $x = 0$. This is usually called the *resonance zone* and the set corresponding with $x = 0$ the *resonance manifold* for reasons that will become clear in the next section. A picture of the phase-plane of the solutions helps us to understand the dynamics (see Fig. 12.1). Starting away from $x = 0$, $x(t)$ increases as a linear function with small modulations. In the resonance zone, the solutions may be captured for all time or pass through the resonance zone.

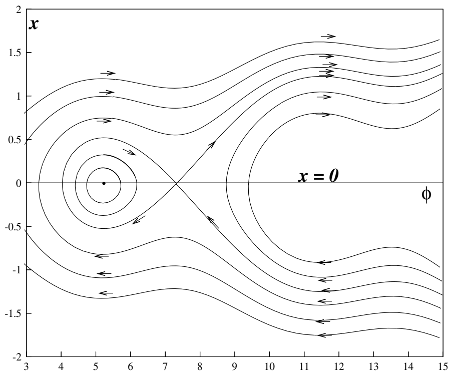

Fig. 12.1. Passage through resonance and capture into resonance in Example 12.4; the resonance zone is located near the resonance manifold $x = 0$.

This resonance zone is really a boundary layer in the sense we discussed in Chapter 4. We can see that by rescaling near $x = 0$: $\xi = x/\delta(\varepsilon)$. Transforming $x, \phi \to \xi, \phi$, the equations become

$$\dot{\xi} = \frac{\varepsilon}{\delta(\varepsilon)}(1 - 2\cos\phi),$$
$$\dot{\phi} = \delta(\varepsilon)\xi.$$

A significant degeneration arises when the terms on the right-hand side are of the same order or

$$\frac{\varepsilon}{\delta(\varepsilon)} = \delta(\varepsilon).$$

We conclude that $\delta(\varepsilon) = \sqrt{\varepsilon}$, which is the size of the resonance zone.

This example of a conservative equation is more typical than it seems. Consider for instance the two much more general scalar equations

$$\begin{aligned}\dot{x} &= \varepsilon f(\phi, x),\\ \dot{\phi} &= g(x),\end{aligned}$$

with f and g smooth functions and $f(\phi, x)$ 2π-periodic in ϕ. Averaging over ϕ is possible outside the zeros of g. Suppose $g(a) = 0$ and rescale

$$\xi = \frac{x - a}{\delta(\varepsilon)}.$$

The equations become

$$\begin{aligned}\delta(\varepsilon)\dot{\xi} &= \varepsilon f(\phi, a + \delta(\varepsilon)\xi),\\ \dot{\phi} &= g(a + \delta(\varepsilon)\xi).\end{aligned}$$

Expanding, we find

$$\begin{aligned}\delta(\varepsilon)\dot{\xi} &= \varepsilon f(\phi, a) + O(\delta(\varepsilon)\varepsilon),\\ \dot{\phi} &= \delta(\varepsilon)g'(a)\xi + O(\delta^2(\varepsilon)).\end{aligned}$$

Again, a significant degeneration arises if we choose $\delta(\varepsilon) = \sqrt{\varepsilon}$, which is the size of the resonance zone. The equations in the resonance zone are to first order

$$\begin{aligned}\dot{\xi} &= \sqrt{\varepsilon} f(\phi, a),\\ \dot{\phi} &= \sqrt{\varepsilon} g'(a)\xi,\end{aligned}$$

which is again a conservative system.

12.2 Averaging over more Angles

When more angles are present, many subtle problems and interesting phenomena arise. We start with a few simple examples.

Example 12.5
Consider first the system

$$\begin{aligned}\dot{x} &= \varepsilon x(\cos\phi_1 + \cos\phi_2),\\ \dot{\phi}_1 &= 1,\\ \dot{\phi}_2 &= 2.\end{aligned}$$

We average the equation for x over ϕ_1 and ϕ_2, which produces the averaged equation $\dot{x}_a = 0$ so that $x_a(t) = x(0)$. The solution $x(t)$ can easily be obtained and reads

$$x(t) = x(0)e^{\varepsilon(\sin(t+\phi_1(0)) + \frac{1}{2}\sin(2t+\phi_2(0)))},$$

so $x_a(t) = x(0)$ is an approximation and this correct answer seems quite natural, as the right-hand sides of the angle equations do not vanish.

We modify the equations slightly in the next example.

Example 12.6
Consider the system

$$\begin{aligned}\dot{x} &= \varepsilon x(\cos\phi_1 + \cos\phi_2 + \cos(2\phi_1 - \phi_2)),\\ \dot{\phi}_1 &= 1,\\ \dot{\phi}_2 &= 2.\end{aligned}$$

We average the equation for x over ϕ_1 and ϕ_2, which produces again the averaged equation $\dot{x}_a = 0$ so that $x_a(t) = x(0)$. Also, the solution $x(t)$ can easily be obtained and reads

$$x(t) = x(0)e^{\varepsilon(\sin(t+\phi_1(0))+\frac{1}{2}\sin(2t+\phi_2(0))+t\cos(2\phi_1(0)-\phi_2(0)))},$$

so $x_a(t) = x(0)$ is in general not an approximation valid on the timescale $1/\varepsilon$. What went wrong?

The right-hand side of the equation for x actually contains three angle combinations, ϕ_1, ϕ_2, and $2\phi_1 - \phi_2$. Averaging should take place over three angles instead of two. Adding formally the (dependent) equation for this third angle, we find that the right-hand side vanishes. The experience from the preceding section tells us that we might expect trouble because of this resonance.

Note that this example is rather extreme in its simplicity. The angles vary with a constant rate so that once we have resonance we have for all values of x a resonance manifold, it fills up the whole x-space. In general, the angles do not vary with a constant rate and behave more as in the following modification.

Example 12.7
Consider the system

$$\begin{aligned}\dot{x} &= \varepsilon x(\cos\phi_1 + \cos\phi_2 + \cos(2\phi_1 - \phi_2)),\\ \dot{\phi}_1 &= x,\\ \dot{\phi}_2 &= 2.\end{aligned}$$

We have still the same three angle combinations in the equation for x as in the preceding example. Resonance can be expected if $\dot{\phi}_1 = 0$, $\dot{\phi}_2 = 0$, or $2\dot{\phi}_1 - \dot{\phi}_2 = 2(x-1) = 0$. This leads to resonance if $x = 0$ and if $x = 1$ with corresponding resonance manifolds. Outside the resonance zones around these manifolds, we can average over the three angles to find the approximation $x_a(t) = x(0)$. How do we visualise the flow?

Outside the resonance manifolds $x = 0$ and $x = 1$, the variable x is nearly constant. To see what happens for instance in the resonance zone near $x = 1$, we put $\psi = 2\phi_1 - \phi_2$ and rescale,

$$\xi = \frac{x-1}{\delta(\varepsilon)}.$$

Introducing this into the equations and expanding, we obtain

$$\begin{aligned}\delta(\varepsilon)\dot{\xi} &= \varepsilon(\cos\phi_1 + \cos\phi_2 + \cos\psi) + O(\varepsilon\delta(\varepsilon)),\\ \dot{\psi} &= 2\delta(\varepsilon)\xi,\\ \dot{\phi}_1 &= 1 + O(\delta(\varepsilon)),\\ \dot{\phi}_2 &= 2.\end{aligned}$$

The three angles are dependent and can be replaced by two angles, for instance ϕ_1 and ψ, but this makes little difference in the outcome. A significant degeneration arises on choosing $\delta(\varepsilon) = \sqrt{\varepsilon}$, which is the size of the resonance zone. We find locally

$$\begin{aligned}\dot{\xi} &= \sqrt{\varepsilon}(\cos\phi_1 + \cos\phi_2 + \cos\psi) + O(\varepsilon),\\ \dot{\psi} &= 2\sqrt{\varepsilon}\xi,\\ \dot{\phi}_1 &= 1 + O(\sqrt{\varepsilon}),\\ \dot{\phi}_2 &= 2.\end{aligned}$$

The system we obtained has two slowly varying variables, ξ and ψ, and two angles with $O(1)$ variation, so we can average over ϕ_1 and ϕ_2 to obtain equations for the first-order approximations ξ_a and ψ_a:

$$\begin{aligned}\dot{\xi}_a &= \sqrt{\varepsilon}\cos\psi_a,\\ \dot{\psi}_a &= 2\sqrt{\varepsilon}\xi_a.\end{aligned}$$

Differentiation yields that ψ_a satisfies the (conservative) pendulum equation

$$\ddot{\psi}_a - 2\varepsilon\cos\psi_a = 0,$$

which describes oscillatory motion in the resonance zone. It is remarkable that we have found a conservative equation at first order, which is of course very sensitive to perturbations. The original system of equations is not conservative, so this suggests that we have to compute a second-order approximation to obtain a structurally stable result.

Before presenting a more general formulation, we consider the phenomenon of *resonance locking*.

Example 12.8
Consider the four-dimensional system

$$\begin{aligned}\dot{x}_1 &= \varepsilon,\\ \dot{x}_2 &= \varepsilon\cos(\phi_1 - \phi_2),\\ \dot{\phi}_1 &= 2x_1,\\ \dot{\phi}_2 &= x_1 + x_2.\end{aligned}$$

There is one angle combination, $\phi_1 - \phi_2$, and we expect resonance if $\dot{\phi}_1 - \dot{\phi}_2 = x_1 - x_2$ vanishes. Outside the resonance zone around the line $x_1 = x_2$ in the x_1, x_2-plane, we can average over the angles to obtain

$$\dot{x}_{1a} = \varepsilon,\ \dot{x}_{2a} = 0,$$

so that $x_{1a} = \varepsilon t + x_1(0), x_{2a} = x_2(0)$. Inside the resonance zone, we can analyse the original equations by putting

$$x = x_1 - x_2,\ \psi = \phi_1 - \phi_2,$$

which leads to the reduced system

$$\begin{aligned} \dot{x} &= \varepsilon(1 - \cos\psi), \\ \dot{\psi} &= x. \end{aligned}$$

Differentiation of the equation for ψ produces

$$\ddot{\psi} + \varepsilon \cos\psi = \varepsilon,$$

which is a forced pendulum equation.

Note that we have resonance locking as the solutions $x = 0$, $\cos\psi = 1$ are equilibrium solutions of the reduced system.

This is typical for many near-integrable Hamiltonian systems where we have in general an infinite number of resonance zones in which resonance locking can take place.

12.2.1 General Formulation of Resonance

We will now give a general formulation for the case of two or more angles. Consider the system

$$\begin{aligned} \dot{x} &= \varepsilon X(\phi, x), \\ \dot{\phi} &= \Omega(x), \end{aligned}$$

with $x \in \mathbb{R}^n, \phi \in T^m$; T^m is the m-dimensional torus described by m angles. Suppose that the vector function X is periodic in the m angles ϕ and that we have the multiple (complex) Fourier expansion

$$X(\phi, x) = \sum_{k_1, \cdots, k_m = -\infty}^{+\infty} c_{k_1, \cdots, k_m}(x) e^{i(k_1\phi_1 + k_2\phi_2 + \cdots k_m\phi_m)}$$

with $(k_1, \cdots, k_m) \in \mathbb{Z}^m$. The resonance manifolds in $\mathbb{R}^n$ (x-space) are determined by the relations

$$k_1\Omega_1(x) + \cdots k_m\Omega_m(x) = 0,$$

assuming that the Fourier coefficient c_k with $k = (k_1, \cdots, k_m)$ does not vanish.

In applications, there are usually order of magnitude variations in the Fourier coefficients so that we can neglect some. If the resonance manifolds do not fill up the whole x-space, we can average *outside* the resonance manifolds to obtain the equation for the approximation $x_a(t)$,

$$\dot{x}_a = \varepsilon c_{0,\cdots,0}(x_a).$$

One can prove that outside the resonance manifolds, assuming that $x(0) = x_a(0)$, we have the estimate

$$x(t) - x_a(t) = O(\varepsilon) \text{ on the timescale } 1/\varepsilon.$$

12.2.2 Nonautonomous Equations

In practice, it happens quite often that time t enters explicitly into the equations. Consider the system

$$\begin{aligned} \dot{x} &= \varepsilon X(\phi, t, x), \\ \dot{\phi} &= \Omega(x), \end{aligned}$$

with the vector function X periodic in t. The scaling to the same period of angles and time is important to make the variations comparable. Suppose that ϕ is m-dimensional; put

$$\phi_{m+1} = t, \;\; \dot{\phi}_{m+1} = 1,$$

and consider averaging over $m+1$ angles.

This procedure is correct, but of course the dependence on t may produce many additional resonances.

Example 12.9
Consider the system

$$\begin{aligned} \dot{x} &= \varepsilon X(\phi_1, t, x), \\ \dot{\phi}_1 &= x, \end{aligned}$$

with $X(\phi_1, t, x) = 2x \sin t \sin \phi_1$. Putting

$$\phi_2 = t, \;\; \dot{\phi}_2 = 1, X(\phi_1, t, x) = x(\cos(\phi_1 - t) - \cos(\phi_1 + t)),$$

we obtain the system with two angles,

$$\begin{aligned} \dot{x} &= \varepsilon x(\cos(\phi_1 - \phi_2) - \cos(\phi_1 + \phi_2)), \\ \dot{\phi}_1 &= x, \\ \dot{\phi}_2 &= 1. \end{aligned}$$

The resonance manifolds correspond with the zeros of the right-hand sides of $\dot{\phi}_1 - \dot{\phi}_2$ and $\dot{\phi}_1 + \dot{\phi}_2$, so we find $x = 1$ and $x = -1$. Outside these resonance zones, we can average over the angles to find $\dot{x}_a = 0$. In the resonance zones, the flow is again described by pendulum equations.
Note that the Fourier expansion of X contains only two terms. If there were an infinite number of terms, we would have resonance relations like

$$k_1 x + k_2 = 0,$$

which would produce resonance manifolds for an infinite number of rational values of x.

12.2.3 Passage through Resonance

We have seen an example of locking into resonance. Interesting phenomena arise when we have passage through resonance; in Example 12.11, we shall see an application.

To start with, we discuss an interesting example of *forced* passage through resonance that was constructed by Arnold (1965).

Example 12.10
A seemingly small variation of an earlier example is the system

$$\begin{aligned} \dot{x}_1 &= \varepsilon, \\ \dot{x}_2 &= \varepsilon \cos(\phi_1 - \phi_2), \\ \dot{\phi}_1 &= x_1 + x_2, \\ \dot{\phi}_2 &= x_2, \end{aligned}$$

with initial values $x_1(0) = -a, x_2(0) = 1, \phi_1(0) = \phi_2(0) = 0$, with a a constant independent of ε. We have one angle combination that can lead to resonance (i.e., if $\dot{\phi}_1 - \dot{\phi}_2 = x_1 = 0$). Integration produces

$$x_1(t) = -a + \varepsilon t,\ \ \phi_1(t) - \phi_2(t) = -at + \frac{1}{2}\varepsilon t^2,$$

and so we have

$$x_2(t) = 1 + \varepsilon \int_0^t \cos\left(-as + \frac{1}{2}\varepsilon s^2\right) ds.$$

If $x_1(0) > 0$ (a negative), the solution does not pass through the resonance zone around $x_1 = 0$. Partial integration produces that $x_2(t) = 1 + O(\varepsilon)$ for all time. If $x_1(0) < 0$ (a positive), we have forced crossing of the resonance zone. In the case $a = 0$, we start in the resonance zone and we can use the well-known integral

$$\int_0^\infty \cos s^2 ds = \frac{1}{2}\sqrt{\frac{\pi}{2}},$$

so if $a = 0$ we find the long-term effect of this crossing by taking the limit $t \to \infty$:

$$\lim_{t\to\infty} x_2(t) = 1 + \frac{1}{2}\sqrt{\pi\varepsilon}.$$

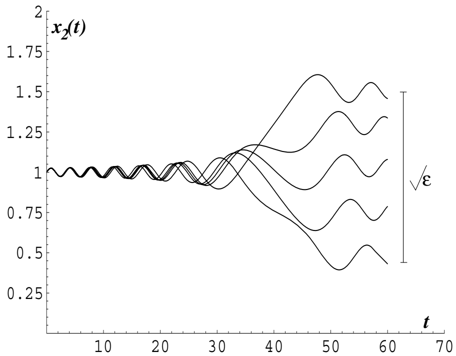

Fig. 12.2. Dispersion of orbits by passage through resonance; the five orbits started at $x_1(0) = -a$ with $a = 2, 2 \pm \varepsilon, 2 \pm \frac{1}{2}\varepsilon$, $\varepsilon = 0.1$.

More generally, we have to calculate or estimate

$$\lim_{t\to\infty} x_2(t) = 1 + \varepsilon \int_0^\infty \cos\left(-as + \frac{1}{2}s^2\right) ds.$$

Transforming

$$s = \sqrt{\frac{2}{\varepsilon}}u + \frac{a}{\varepsilon},$$

we find

$$\varepsilon \int_0^\infty \cos\left(-as + \frac{1}{2}s^2\right) ds = \sqrt{2\varepsilon}\int_{-\frac{a}{\sqrt{2\varepsilon}}}^\infty \cos\left(-\frac{a^2}{2\varepsilon} + u^2\right) du,$$

which can be split into

$$\sqrt{2\varepsilon}\left(\cos\left(\frac{a^2}{2\varepsilon}\right)\int_{-\frac{a}{\sqrt{2\varepsilon}}}^\infty \cos u^2 du + \sin\left(\frac{a^2}{2\varepsilon}\right)\int_{-\frac{a}{\sqrt{2\varepsilon}}}^\infty \sin u^2 du\right).$$

The two integrals equal $\sqrt{(\pi/2)} + o(1)$, so that we have an estimate for the long-term effect of passing through resonance,

$$\lim_{t\to\infty} x_2(t) = 1 + \sqrt{\pi\varepsilon}\left(\cos\left(\frac{a^2}{2\varepsilon}\right) + \sin\left(\frac{a^2}{2\varepsilon}\right) + o(1)\right).$$

This shows that the effect of passing through resonance is $O(\sqrt{\varepsilon})$ and remarkably that the solution displays sensitive dependence on the initial condition. Small changes of a produce relatively large changes of the solution. This "dispersion" of orbits is illustrated in Fig. 12.2.

Finally, we shall briefly discuss an application in mechanics that displays both passage through and (undesirable) capture into resonance.

Example 12.11
Consider a spring that can move in the vertical x direction on which a rotating wheel is mounted; the rotation angle is ϕ. The wheel has a small mass fixed on the edge that makes it slightly eccentric, a flywheel. The vertical displacement x and the rotation ϕ are determined by the equations

$$\ddot{x} + x = \varepsilon(-x^3 - \dot{x} + \dot{\phi}^2 \cos\phi) + O(\varepsilon^2),$$
$$\ddot{\phi} = \varepsilon\left(\frac{1}{4}(2 - \dot{\phi}) + (1 - x)\sin\phi\right) + O(\varepsilon^2).$$

See Evan-Ewanowski (1976) for the equations; we have added an appropriate scaling, assuming that the friction, the nonlinear restoring force, the eccentric mass, and several other forces are small.

To obtain a standard form suitable for averaging, we transform

$$x = r\sin\phi_2,\ \dot{x} = r\cos\phi_2,\ \phi = \phi_1,\ \dot{\phi}_1 = \Omega,$$

with $r > 0, \Omega > 0$. This introduces two angles and two slowly varying quantities:

$$\begin{aligned}
\dot{r} &= \varepsilon\cos\phi_2(-r^3\sin^3\phi_2 - r\cos\phi_2 + \Omega^2\cos\phi_1),\\
\dot{\Omega} &= \varepsilon\left(\frac{1}{4}(2-\Omega) + \sin\phi_1 - r\sin\phi_1\sin\phi_2\right),\\
\dot{\phi}_1 &= \Omega,\\
\dot{\phi}_2 &= 1 + \varepsilon\left(r^2\sin^4\phi_2 + \frac{1}{2}\sin 2\phi_2 - \frac{\Omega^2}{r}\cos\phi_1\sin\phi_2\right).
\end{aligned}$$

The $O(\varepsilon^2)$ terms have been omitted. Resonance zones exist if

$$m\Omega + n = 0,\ m, n \in \mathbb{Z}.$$

In the equation for r and Ω to $O(\varepsilon)$, the angles are $\phi_1, \phi_2, \phi_1 + \phi_2, \phi_1 - \phi_2$. As ϕ_1 and ϕ_2 are monotonically increasing, the only resonance zone that can arise is when $\phi_1 - \phi_2 = 0$, which determines the resonance manifold $\Omega = 1$. Outside the resonance zone, a neighbourhood of $\Omega = 1$, we average over the angles to find the approximations r_a and Ω_a given by

$$\dot{r}_a = -\frac{1}{2}\varepsilon r_a,$$
$$\dot{\Omega}_a = \frac{1}{4}\varepsilon(2 - \Omega_a).$$

This is already an interesting result. Outside the resonance zone, $r(t) = r_a(t) + O(\varepsilon)$ will decrease exponentially with time; on the other hand, $\Omega(t)$ will tend to the value 2. If we start with $\Omega(0) < 1$, $\Omega(t)$ will after some time enter the resonance zone around $\Omega = 1$. How does this affect the dynamics? Will the system pass in some way through resonance or will it stay in the resonance zone, resulting in vertical oscillations that are undesirable for a mounted flywheel.

The way to answer these questions is to analyse what is going on in the resonance zone and find out whether there are attractors present. Following the analysis of localising into the resonance zone as before, we introduce the resonant combination angle $\psi = \phi_1 - \phi_2$ and the local variable

$$\omega = \frac{\Omega - 1}{\sqrt{\varepsilon}}.$$

Transforming the equations for r, Ω and the angles, the leading terms are $O(\sqrt{\varepsilon})$; we find

$$\begin{aligned}
\dot{r} &= \varepsilon \cdots , \\
\dot{\omega} &= \sqrt{\varepsilon}\left(\frac{1}{4} + \sin\phi_1 - \frac{1}{2}r\cos\psi + \frac{1}{2}r\cos(2\phi_1 - \psi)\right) + \varepsilon \cdots , \\
\dot{\psi} &= \sqrt{\varepsilon}\omega + \varepsilon \cdots , \\
\dot{\phi}_1 &= 1 + \sqrt{\varepsilon}\omega.
\end{aligned}$$

We can average over the remaining angle ϕ_1; as the equation for r starts with $O(\varepsilon)$ terms, we have in the resonance zone that $r(t) = r(0) + O(\sqrt{\varepsilon})$. The equations for the approximations of ω and ψ are

$$\begin{aligned}
\dot{\omega}_a &= \sqrt{\varepsilon}\left(\frac{1}{4} - \frac{1}{2}r\cos\psi\right), \\
\dot{\psi} &= \sqrt{\varepsilon}\omega.
\end{aligned}$$

By differentiation of the equation for ψ_a, we can write this as the pendulum equation

$$\ddot{\psi}_a + \frac{1}{2}\varepsilon r(0)\cos\psi_a = \frac{1}{4}\varepsilon.$$

The timescale of the dynamics is clearly $\sqrt{\varepsilon}t$; there are two equilibria, one a centre point and the other a saddle. They correspond with periodic solutions of the original system. The saddle is definitely unstable, and for the centre point we have to perform higher-order averaging, to $O(\varepsilon)$, to determine the stability. This analysis was carried out by Van den Broek (1988); see also Van den Broek and Verhulst (1987). The result is that by adding $O(\varepsilon)$ terms, the centre point in the resonance zone becomes an attracting focus so that the corresponding periodic solution is stable.

The implication is that for certain initial values the oscillator-flywheel might pass into resonance and stay there. Van den Broek (1988) identified three sets of initial values leading to capture into resonance.

1. **Remark**
   An extension of the theory of averaging over angles is possible for systems of the form

$$\dot{x} = \varepsilon X(\phi, x),$$
$$\dot{\phi} = \Omega(\phi, x).$$

   This generalisation complicates the calculations; see Section 5.4 in Sanders and Verhulst (1985).
2. **Remark**
   In the examples studied here, we obtained pendulum equations describing the flow in the resonance zones of the respective cases. This was observed by many authors in examples. In Section 11.7 of Verhulst (2000), it is shown that a first-order (in ε) computation in a resonance zone always leads to a conservative equation - often a pendulum equation or system of pendulum equations - describing the flow. This is remarkable, as the original system need not be conservative at all and the first-order result will probably change qualitatively under perturbation. The result stresses again the importance of second-order calculations in these cases.

12.3 Invariant Manifolds

An important problem is to determine invariant manifolds such as tori or cylinders in nonlinear equations. Consider a system such as

$$\dot{x} = f(x) + \varepsilon R(t, x, \varepsilon).$$

Suppose for instance that we have found an isolated torus T_a by first-order averaging. Does this manifold persist, slightly deformed as a torus T, when considering the original equation? Note that the original equation can be seen as a perturbation of the averaged equation, and the question can then be rephrased as the question of persistence of the torus T_a under perturbation. If the torus in the averaged equation is *normally hyperbolic*, the answer is affirmative. Normally hyperbolic means, loosely speaking, that the strength of the flow along the manifold is weaker than the rate of attraction to the manifold. We have used such results in Chapter 9 for Tikhonov-Fenichel problems. In many applications, however, the approximate manifold that one obtains is hyperbolic but not normally hyperbolic. In the Hamiltonian case, the tori arise in families and they will not even be hyperbolic.

We will look at different scenarios for the emergence of tori in some examples. A torus is generated by various independent rotational motions - at least two - and we shall find different timescales characterising these rotations.

12.3.1 Tori in the Dissipative Case

First, we look at cases where the branching off of tori is similar to the emergence of periodic solutions in the examples we have seen before. The theory of such questions was considered extensively by Bogoliubov and Mitropolsky (1961) and uses basically continuation of quasiperiodic motion under perturbations; for a summary and other references, see also Bogoliubov and Mitropolsky (1963). Another survey and new results can be found in Hale (1969); see the references therein.

There are many interesting open problems in this field, as the bifurcation theory of invariant manifolds is clearly even richer than for equilibria or periodic solutions. We present a few illustrative examples.

Example 12.12
Consider the system

$$\ddot{x} + x = \varepsilon\left(2x + 2\dot{x} - \frac{8}{3}\dot{x}^3 + y^2x^2 + \dot{y}^2x^2\right) + \varepsilon^2 R_1(x, y),$$
$$\ddot{y} + \omega^2 y = \varepsilon(\dot{y} - \dot{y}^3 + x^2y^2 + \dot{x}^2y^2) + \varepsilon^2 R_2(x, y),$$

where R_1 and R_2 are smooth functions. Introducing amplitude-phase coordinates by $x = r_1\cos(t + \psi_1), \dot{x} = -r_1\sin(t + \psi_1), y = r_2\cos(\omega t + \psi_2), \dot{y} = -\omega r_2\sin(\omega t + \psi_1)$, and after first-order averaging, we find, omitting the subscripts a, the system

$$\dot{r}_1 = \varepsilon r_1(1 - r_1^2), \dot{\psi}_1 = -\varepsilon,$$
$$\dot{r}_2 = \varepsilon\frac{r_2}{2}\left(1 - \frac{3}{4}r_2^2\right), \dot{\psi}_2 = 0.$$

The averaged equations contain a torus T_a in phase-space described by

$$x_a(t) = \cos(t - \varepsilon t + \psi_1(0)), \dot{x}_a(t) = -\sin(t - \varepsilon t + \psi_1(0)),$$
$$y_a(t) = \frac{2}{3}\sqrt{3}\cos(\omega t + \psi_2(0)), \dot{y}_a(t) = -\frac{2\omega}{3}\sqrt{3}\sin(\omega t + \psi_2(0)).$$

From linearisation of the averaged equations, it is clear that the torus is attracting: it is hyperbolic but not normally hyperbolic, as the motion along the torus has $O(1)$ speed and the attraction rate is $O(\varepsilon)$. If the ratio of $1 - \varepsilon$ and ω is rational, the torus T_a is filled up with periodic solutions. If the ratio is irrational, we have a quasiperiodic (two-frequency) flow over the torus. Remarkably enough, the theorems in the literature cited above tell us that in the original equations a torus T exists in an $O(\varepsilon)$ neighbourhood of T_a with the same stability properties. The torus is two-dimensional and the timescales of rotation are in both directions $O(1)$.

The next example was formulated as an exercise by Hale (1969). It is a rich problem and we cannot discuss all of its aspects.

Example 12.13
Consider the system

$$\ddot{x} + x = \varepsilon(1 - x^2 - ay^2)\dot{x},$$
$$\ddot{y} + \omega^2 y = \varepsilon(1 - y^2 - \alpha x^2)\dot{y},$$

with ε-independent positive constants a, α, ω. Using the same amplitude-phase transformation as in the preceding example, we find the slowly varying system

$$\begin{aligned}
\dot{r}_1 &= \varepsilon r_1 \sin(t+\psi_1)(1 - r_1^2\cos^2(t+\psi_1) - ar_2^2\cos^2(\omega t+\psi_2))\sin(t+\psi_1),\\
\dot{\psi}_1 &= \varepsilon \cos(t+\psi_1)(1 - r_1^2\cos^2(t+\psi_1) - ar_2^2\cos^2(\omega t+\psi_2))\sin(t+\psi_1),\\
\dot{r}_2 &= \varepsilon r_2 \sin(\omega t+\psi_2)(1 - r_2^2\cos^2(\omega t+\psi_2) - \alpha r_1^2\cos^2(t+\psi_1))\sin(\omega t+\psi_2),\\
\dot{\psi}_2 &= \varepsilon \cos(\omega t+\psi_2)(1 - r_2^2\cos^2(\omega t+\psi_2) - \alpha r_1^2\cos^2(t+\psi_1))\sin(\omega t+\psi_2).
\end{aligned}$$

First-order averaging yields different results in two cases, $\omega \neq 1$ and $\omega = 1$. However, in all cases we have the following periodic solutions that are also present as solutions of the original equations.

Normal modes
Putting $r_1 = 0, r_2 = 2$ produces a normal mode periodic solution P_2 in the $y, \dot{y}$ coordinate plane. In the same way, we obtain a normal mode periodic solution P_1 in the $x, \dot{x}$ coordinate plane by putting $r_2 = 0, r_1 = 2$. These normal modes in the coordinate planes also exist in the original system. Their stability is studied by linearisation of the averaged equations.

$\omega \neq 1$
The averaged equations are (we omit again the subscript a)

$$\begin{aligned}
\dot{r}_1 &= \frac{\varepsilon}{2} r_1 \left(1 - \frac{1}{4}r_1^2 - \frac{1}{2}ar_2^2\right),\\
\dot{\psi}_1 &= 0,\\
\dot{r}_2 &= \frac{\varepsilon}{2} r_2 \left(1 - \frac{1}{4}r_2^2 - \frac{1}{2}\alpha r_1^2\right),\\
\dot{\psi}_2 &= 0.
\end{aligned}$$

Linearisation around the normal modes produces matrices with many zeros. In the case $r_1 = 0, r_2 = 2$ (P_2), we find for the derivative in the $y, \dot{y}$-plane $-\varepsilon$, which means attraction in this plane; in the $x, \dot{x}$-plane, we find for the derivative $\frac{1}{2}\varepsilon(1-2a)$, which means attraction if $a > \frac{1}{2}$ and repulsion if $a < \frac{1}{2}$. In the case of repulsion, we have instability of the $r_1 = 0, r_2 = 2$ normal mode. In the same way, we find instability of the $r_2 = 0, r_1 = 2$ normal mode P_1 in the $x, \dot{x}$-plane if $\alpha < \frac{1}{2}$.

We will now study the flow outside the coordinate planes. Stationary solutions outside the normal modes can be found if simultaneously

$$1 - \frac{1}{4}r_1^2 - \frac{1}{2}ar_2^2 = 0,\ 1 - \frac{1}{4}r_2^2 - \frac{1}{2}\alpha r_1^2 = 0.$$

These relations correspond with quadrics in the r_1, r_2-plane. They intersect,

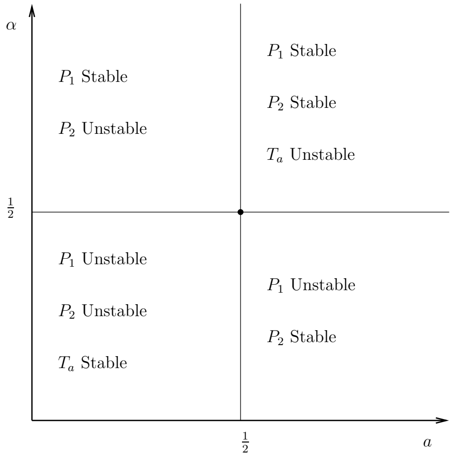

Fig. 12.3. Diagram of invariant manifolds, periodic solutions, and tori in Example 12.13; the point $a = \frac{1}{2}, \alpha = \frac{1}{2}$ is exceptional.

producing one solution, in the cases $a < \frac{1}{2}, \alpha < \frac{1}{2}$ and $a > \frac{1}{2}, \alpha > \frac{1}{2}$. In these two cases, the stationary solutions of the averaged equations correspond with a torus T_a. By linearisation of the averaged equations, we can establish the instability of the torus if $a > \frac{1}{2}, \alpha > \frac{1}{2}$. Stability cannot be deduced from the averaged system in these coordinates, as the matrix is singular. However, it is interesting to write in this case the original system in amplitude-angle coordinates and perform averaging over two angles.

Putting $x = r_1 \sin\phi_1, \dot{x} = r_1 \cos\phi_1, y = r_2 \sin(\omega\phi_2), \dot{y} = r_2\omega\cos(\omega\phi_2)$, we find when averaging over ϕ_1, ϕ_2

$$\dot{r}_1 = \frac{\varepsilon}{2}r_1\left(1 - \frac{1}{4}r_1^2 - \frac{1}{2}ar_2^2\right), \dot{\phi}_1 = 1 + O(\varepsilon),$$
$$\dot{r}_2 = \frac{\varepsilon}{2}r_2\left(1 - \frac{1}{4}r_2^2 - \frac{1}{2}\alpha r_1^2\right), \dot{\phi}_2 = 1 + O(\varepsilon).$$

For the system in amplitude-angle coordinates, we can apply the theory cited above with the conclusion that the torus T_a, when it exists, corresponds with a torus T of the original equations. The torus is stable if $a < \frac{1}{2}, \alpha < \frac{1}{2}$ and unstable if $a > \frac{1}{2}, \alpha > \frac{1}{2}$; see Fig. 12.3. One can see immediately that a critical case is $a = \alpha = \frac{1}{2}$. In this case, the averaged equations contain an invariant sphere $r_1^2 + r_2^2 = 4$ in four-dimensional phase-space. For the question of persistence of this invariant sphere (slightly deformed) in the original equations, first-order averaging is sufficient. The sphere is a first-order approximation of a centre manifold. For the flow *on* the sphere, we need higher-order approximations. Outside this special point, we can cross the lines $a = \frac{1}{2}, \alpha = \frac{1}{2}$ to pass from a phase-space with two periodic solutions to a phase-space of two periodic solutions and a torus. When crossing these lines, one of the normal modes changes stability and produces the torus by a so-called branching bifurcation. The results are summarised in the diagram of Fig. 12.3. In this example, the torus again is two-dimensional and the timescales of rotation are in both directions $O(1)$.

$\omega = 1$
We consider now the special resonance case when the basic frequencies are equal. The averaged equations are

$$\begin{aligned}
\dot{r}_1 &= \frac{\varepsilon}{2} r_1 \left(1 - \frac{1}{4} r_1^2 - \frac{1}{2} a r_2^2 + \frac{1}{4} a r_2^2 \cos 2(\psi_1 - \psi_2)\right), \\
\dot{\psi}_1 &= -\frac{\varepsilon}{8} a r_2^2 \sin 2(\psi_1 - \psi_2), \\
\dot{r}_2 &= \frac{\varepsilon}{2} r_2 \left(1 - \frac{1}{4} r_2^2 - \frac{1}{2} \alpha r_1^2 + \frac{1}{4} \alpha r_1^2 \cos 2(\psi_1 - \psi_2)\right), \\
\dot{\psi}_2 &= \frac{\varepsilon}{8} \alpha r_1^2 \sin 2(\psi_1 - \psi_2).
\end{aligned}$$

Using the combination angle $\chi = 2(\psi_1 - \psi_2)$, the equations become

$$\begin{aligned}
\dot{r}_1 &= \frac{\varepsilon}{2} r_1 \left(1 - \frac{1}{4} r_1^2 - \frac{1}{2} a r_2^2 + \frac{1}{4} a r_2^2 \cos \chi\right), \\
\dot{r}_2 &= \frac{\varepsilon}{2} r_2 \left(1 - \frac{1}{4} r_2^2 - \frac{1}{2} \alpha r_1^2 + \frac{1}{4} \alpha r_1^2 \cos \chi\right), \\
\dot{\chi} &= -\frac{\varepsilon}{4} (\alpha r_1^2 + a r_2^2) \sin \chi.
\end{aligned}$$

Apart from the normal modes P_1 and P_2, we find phase-locked periodic solutions in a general position by putting $\sin \chi = 0$. Introducing this condition into the equations for the amplitudes, we find that for these solutions to exist we have

$$\begin{aligned}
1 - \frac{1}{4} r_1^2 - \frac{1}{2} a r_2^2 + \pm \frac{1}{4} a r_2^2 &= 0, \\
1 - \frac{1}{4} r_2^2 - \frac{1}{2} \alpha r_1^2 + \pm \frac{1}{4} \alpha r_1^2 &= 0.
\end{aligned}$$

The analysis of these equations and the corresponding stability is a lot of work but straightforward and is left to the reader.

As stated in the introduction to this section, when we start off with a normally hyperbolic torus, small perturbations will only deform the torus. A simple example follows.

Example 12.14

$$\ddot{x} + x = \mu(1 - x^2)\dot{x} + \varepsilon f(x, y),$$
$$\ddot{y} + \omega^2 y = \mu(1 - y^2)\dot{y} + \varepsilon g(x, y),$$

with ε-independent positive constant ω, μ a fixed large, positive number, and smooth perturbations f, g. Omitting the perturbations f, g, we have two normally hyperbolic relaxation oscillations. If ω is irrational, the combined oscillations attract to a torus filled with quasiperiodic motion. Adding the perturbations f, g cannot destroy this torus but only deforms it. Also, in this example the torus is two-dimensional but the timescales of rotation are in both directions determined by the timescales of relaxation oscillation (see Grasman, 1987) and so $O(1/\mu)$.

12.3.2 The Neimark-Sacker Bifurcation

Another important scenario for creating a torus arises from the Neimark-Sacker bifurcation. Suppose that we have obtained an averaged equation $\dot{x} = \varepsilon f(x, a)$ with dimension 3 or higher by variation of constants and subsequent averaging; a is a parameter or a set of parameters. We have discussed before that if this equation contains a hyperbolic critical point, the original equation contains a periodic solution. The first-order approximation of this periodic solution is characterised by the timescales t and εt.

Suppose now that by varying the parameter a a pair of eigenvalues of the critical point becomes purely imaginary. For this value of a, the averaged equation undergoes a Hopf bifurcation, producing a periodic solution of the averaged equation; the typical timescale of this periodic solution is εt, so the period will be $O(1/\varepsilon)$. As it branches off an existing periodic solution in the original equation, it will produce a torus, and the bifurcation has a different name: the Neimark-Sacker bifurcation. The result will be a two-dimensional torus that contains two-frequency oscillations, one on a timescale of order 1 and the other with timescale $O(1/\varepsilon)$. A typical example runs as follows.

Example 12.15
A special case of a system studied by Bakri et al. (2004) is

$$\ddot{x} + \varepsilon\kappa\dot{x} + (1 + \varepsilon\cos 2t)x + \varepsilon xy = 0,$$
$$\ddot{y} + \varepsilon\dot{y} + 4(1 + \varepsilon)y - \varepsilon x^2 = 0.$$

This is a system with parametric excitation and nonlinear coupling; κ is a positive damping coefficient that is independent of ε. Away from the coordinate planes, we may use amplitude-phase variables with $x = r_1\cos(t + \psi_1), \dot{x} = -r_1\sin(t + \psi_1), y = r_2\cos(2t + \psi_2), \dot{y} = -2r_2\sin(2t + \psi_1)$; after first-order averaging, we find, omitting the subscripts a, the system

$$\dot{r}_1 = \varepsilon r_1 \left(\frac{r_2}{4} \sin(2\psi_1 - \psi_2) + \frac{1}{4} \sin 2\psi_1 - \frac{1}{2}\kappa \right),$$
$$\dot{\psi}_1 = \varepsilon \left(\frac{r_2}{4} \cos(2\psi_1 - \psi_2) + \frac{1}{4} \cos 2\psi_1 \right),$$
$$\dot{r}_2 = \varepsilon \frac{r_2}{2} \left(\frac{r_1^2}{4r_2} \sin(2\psi_1 - \psi_2) - 1 \right),$$
$$\dot{\psi}_2 = \frac{\varepsilon}{2} \left(-\frac{r_1^2}{4r_2} \cos(2\psi_1 - \psi_2) + 2 \right).$$

Putting the right-hand sides equal to zero produces a nontrivial critical point corresponding with a periodic solution of the system for the amplitudes and phases and so a quasiperiodic solution of the original coupled system in x and y. We find for this critical point the relations

$$r_1^2 = 4\sqrt{5}r_2, \cos(2\psi_1 - \psi_2) = \frac{2}{\sqrt{5}}, \sin(2\psi_1 - \psi_2) = \frac{1}{\sqrt{5}}, r_1 = \sqrt{2\kappa + \sqrt{5 - 16\kappa^2}}.$$

This periodic solution exists if the damping coefficient is not too large: $0 \leq \kappa < \frac{\sqrt{5}}{4}$. Linearisation of the averaged equations at the critical point while using these relations produces the matrix

$$A = \begin{pmatrix} 0 & 0 & \frac{r_1}{4\sqrt{5}} & -\frac{r_1^3}{40} \\ 0 & -\kappa & \frac{1}{2\sqrt{5}} & \frac{r_1^2}{80} \\ \frac{r_1}{4\sqrt{5}} & \frac{r_1^2}{2\sqrt{5}} & -\frac{1}{2} & -\frac{r_1^2}{4\sqrt{5}} \\ -\frac{2}{r_1} & 1 & \frac{4\sqrt{5}}{r_1^2} & -\frac{1}{2} \end{pmatrix}.$$

Another condition for the existence of the periodic solution is that the critical point is hyperbolic (i.e., the eigenvalues of the matrix A have no real part zero). It is possible to express the eigenvalues explicitly in terms of κ by using a software package such as MATHEMATICA. However, the expressions are cumbersome. Hyperbolicity is the case if we start with values of κ just below $\frac{\sqrt{5}}{4} = 0.559$. Diminishing κ, we find when $\kappa = 0.546$ that the real part of two eigenvalues vanishes. This value corresponds with a Hopf bifurcation that produces a nonconstant periodic solution of the averaged equations. This in turn corresponds with a torus in the orginal equations (in x and y) by a Neimark-Sacker bifurcation. As stated before, the result will be a two-dimensional torus that contains two-frequency oscillations, one frequency on a timescale of order 1 and the other with timescale $O(1/\varepsilon)$.

12.3.3 Invariant Tori in the Hamiltonian Case

Integrability of a Hamiltonian system means, loosely speaking, that the system has at least as many independent first integrals as the number of degrees of freedom. As each degree of freedom corresponds with one position and one momentum, n degrees of freedom will mean a $2n$-dimensional system of differential equations of motion. With n independent integrals, the special structure of Hamiltonian systems will then invoke a complete foliation of phase-space into invariant manifolds.

Most Hamiltonian systems are nonintegrable, but in practice most are near an integrable system. A fundamental question is then how many of these invariant manifolds survive the nonintegrable perturbation. This question was solved around 1960 in the celebrated KAM theorem, which tells us that near a stable equilibrium point under rather general conditions, an infinite subset of invariant tori will survive. There is now extensive literature on KAM theory; for introductions, see Arnold (1978) or Verhulst (2000). This is an extensive subject and we restrict ourselves to an example here.

Example 12.16
A classical mechanical example is the elastic pendulum. Consider a spring that can both oscillate in the vertical z direction and swing like a pendulum with angular deflection ϕ, where the corresponding momenta are p_z and p_ϕ. The model is discussed in many places, see for instance Van der Burgh (1975) or Tuwankotta and Verhulst (2000). It is shown there that near the vertical rest position, the Hamiltonian can be expanded as $H = H_0 + H_2 + \varepsilon H_3 + \varepsilon^2 H_4 + O(\varepsilon^3)$ with H_0 a constant; ε measures the deflection from the rest position. We have

$$\begin{aligned} H_2 &= \frac{1}{2}\omega_z\left(z^2 + p_z^2\right) + \frac{1}{2}\omega_\phi\left(\phi^2 + p_\phi^2\right), \\ H_3 &= \frac{\omega_\phi}{\sqrt{\sigma\omega_z}}\left(\frac{1}{2}z\phi^2 - zp_\phi^2\right), \\ H_4 &= \frac{1}{\sigma}\left(\frac{3}{2}\frac{\omega_\phi}{\omega_z}z^2 p_\phi^2 - \frac{1}{24}\phi^4\right), \end{aligned}$$

with positive frequencies $\omega_z, \omega_\phi; \sigma$ is a positive constant depending on the mass and the length of the pendulum. As expected from the physical setup, the relatively few terms in the Hamiltonian are symmetric in the second degree of freedom and also in p_z. Due to the physical restrictions, we will have for the frequency ratio $\omega_z/\omega_\phi > 1$.

Fixing $\omega_z/\omega_\phi \neq 2$, we find that by averaging to first order all terms vanish. For these frequency ratios, interesting phenomena take place either on a longer timescale or on a smaller scale with respect to ε. The important lower-order resonance is the $2:1$ resonance, which has been intensively studied. This resonance is the one with resonant terms of the lowest degree.
The case $\omega_z = 2, \omega_\phi = 1$.
The equations of motion derived from the Hamiltonian can, after rescaling of parameters, be written as

$$\ddot{z} + 4z = \varepsilon \left(\dot{\phi}^2 - \frac{1}{2}\phi^2 \right) + O(\varepsilon^2),$$
$$\ddot{\phi} + \phi = \varepsilon(z\phi - 2\dot{z}\dot{\phi}) + O(\varepsilon^2).$$

Introduce the transformation $z = r_1 \cos(2t + \psi_1), \dot{z} = -2r_1 \sin(2t + \psi_1), \phi = r_2 \cos(t + \psi_2), \dot{\phi} = -r_2 \sin(t + \psi_2)$.

An objection against these amplitude-phase transformations is that they are not canonical (i.e., they do not preserve the Hamiltonian structure of the equations of motion). However, as we know, they yield asymptotically correct results; also, the averaging process turns out to conserve the energy. We omit the standard form and give directly the first-order averaging result, leaving out the approximation index a:

$$\dot{r}_1 = -\varepsilon \frac{3}{16} r_2^2 \sin \chi,$$
$$\dot{\phi}_1 = \varepsilon \frac{3r_2^2}{16r_1} \cos \chi,$$
$$\dot{r}_2 = \varepsilon \frac{3}{4} r_1 r_2 \sin \chi,$$
$$\dot{\phi}_2 = \varepsilon \frac{3r_1}{4} \cos \chi,$$

with $\chi = 2\psi_2 - \psi_1$. Actually, because of the presence of the combination angle χ, the system can be reduced to three equations. It is easy to find two integrals of this system. First is the approximate energy integral

$$4r_1^2 + r_2^2 = 2E,$$

with E indicating the constant, initial energy. The second integral is cubic and reads

$$z\phi^2 \cos \chi = I,$$

with I a constant determined by the initial conditions.

Note that the approximate energy integral represents a family of ellipsoids in four-dimensional phase-space around stable equilibrium. For each value of the energy, the second integral induces a foliation of the energy manifold into invariant tori. In this approximation, the foliation is a continuum, but for the original system the KAM theorem guarantees the existence of an infinite number of invariant tori with gaps in between. The gaps cannot be "seen" in a first-order approximation, and remarkably enough, they cannot be found at any algebraic order of approximation.The gaps turn out to be exponentially small (i.e., of size $\varepsilon^a \exp(-b/\varepsilon^c)$ with suitable constants a, b, c). This analysis is typical for time-independent Hamiltonian systems with two degrees of freedom and to some extent for more degrees of freedom. A description and references can be found in Lochak et al. (2003).

In this case of an infinite set of invariant tori, we have basically two timescales. Embedded on the energy manifold are periodic solutions with period $O(1)$. Around the stable periodic solutions, the tori are nested, with the periodic solutions as guiding centers. In the direction of the periodic solution, the timescale is $O(1)$, and in the other direction(s) it is $O(1/\varepsilon)$.

The case of two degrees of freedom is easiest to visualise. The energy manifold is three-dimensional, and a section perpendicular to a stable periodic solution is two-dimensional. In this section, called a Poincaré section, the periodic solution shows up in a few points where it subsequently hits the section. Around these fixed points, we find closed curves corresponding with the tori. Orbits moving on such a torus hit a particular closed curve recurrently with circulation time on the closed curve $O(1/\varepsilon)$.

12.4 Adiabatic Invariants

In Example 12.2, we considered a linear oscillator with slowly varying (prescribed) frequency:

$$\ddot{x} + \omega^2(\varepsilon t)x = 0.$$

Putting $\tau = \varepsilon t$ and assuming that $0 < a < \omega(\tau) < b$ (with a, b constants independent of ε), we found

$$r_a(\tau) = \frac{r(0)\sqrt{\omega(0)}}{\sqrt{\omega(\tau)}},$$

so that the quantity $r_a(\varepsilon t)\sqrt{\omega(\varepsilon t)}$ is conserved in time with accuracy $O(\varepsilon)$ on the timescale $1/\varepsilon$. Such a quantity we call an adiabatic invariant for the equation. More generally, consider the n-dimensional equation $\dot{x} = f(x, \varepsilon t, \varepsilon)$ and suppose that we have found a function $I(x, \varepsilon t)$ with the property

$$I(x, \varepsilon t) = I(x(0), 0) + o(1)$$

on a timescale that tends to infinity as ε tends to zero. In this case, we will call $I(x, \varepsilon t)$ an *adiabatic invariant* of the equation.

Note that this concept is a generalisation of the concept of a “first integral” of a system. The equation $\dot{x} = f(x, \varepsilon t, \varepsilon)$ may not have a first integral but, to a certain approximation, it behaves as if it has the adiabatic invariant as a first integral.

In a number of cases in the literature, the concept of an adiabatic invariant is also used for equations of the type $\dot{x} = f(x, \varepsilon)$. However, in such a case we prefer the term “asymptotic integral”.

In some cases, the simple trick using $\tau = \varepsilon t$ as a dependent variable produces results. Consider the following nearly trivial example.

Example 12.17
Replace the equation

$$\ddot{x} + x = \varepsilon a(\varepsilon t)x^3,$$

with $a(\varepsilon t)$ a smooth function, by the system

$$\ddot{x} + x = \varepsilon a(\tau)x^3, \ \dot{\tau} = \varepsilon.$$

First-order averaging in amplitude-angle coordinates r, ϕ as in Example 12.1 for the Duffing equation can easily be carried out. Supposing that $r(0) = r_0$, we have the approximation $r(t) = r_0 + O(\varepsilon)$ valid on the timescale $1/\varepsilon$. The implication is that the quadratic integral $x^2(t) + \dot{x}^2(t) = r_0^2$ is conserved to $O(\varepsilon)$ on the timescale $1/\varepsilon$. This is a simple but nontrivial adiabatic invariant for the system.

A much more difficult example arises when the perturbation term depends on εt but is not small. Consider for instance the equation

$$\ddot{x} + x = a(\varepsilon t)x^2.$$

Huveneers and Verhulst (1997) studied this equation in the case where $a(0) = 1$ and the smooth function $a(\varepsilon t)$ vanishes as $t \to \infty$. In this case, there exist bounded and unbounded solutions and the analysis leans heavily on averaging over elliptic functions that are the solutions of the equation

$$\ddot{x} + x = x^2.$$

The analysis is too technical to include here.

The example of the Duffing equation is typical for the treatment of Hamiltonian systems with one degree of freedom and slowly varying coefficients and can be found in many texts on classical mechanics; see for instance Arnold (1978). With additional conditions, it is sometimes possible to extend the timescale of validity of the adiabatic invariant beyond $1/\varepsilon$.

Such examples can also be extended to more degrees of freedom, but the analysis becomes much more subtle. The extension of the timescale beyond $1/\varepsilon$ is then not easy to reach. Fascinating problems arise when the slowly varying coefficient leads the system through a bifurcation value. Unfortunately, this topic is beyond the scope of this book, but we give some references in our guide to the literature at the end of this chapter.

12.5 Second-Order Periodic Averaging

To calculate a second-order or even higher-order approximation, generally takes a much larger effort than in first order. One reason for this extra effort could be to obtain an approximation with higher precision but a more fundamental motivation derives from the fact that in quite a number of cases essential qualitative features are not described by first order. Let us consider a few simple examples. For the Van der Pol equation (Example 11.3)

$$\ddot{x} + x = \varepsilon \dot{x}(1 - x^2),$$

we have obtained a first-order approximation of the unique periodic solution. Suppose we add damping of a slightly larger magnitude. A model equation would be

$$\ddot{x} + \varepsilon\mu\dot{x} + x = \varepsilon^2 \dot{x}(1 - x^2),$$

where the damping constant μ is positive. At first order, the solutions are damped, but do we recover a periodic solution at second order? In this case the answer is "no," as we know from qualitative information about this particular type of equation; see for instance Verhulst (2000). In most research problems, we do not have much a priori knowledge. A slightly less trivial modification arises when at first order we have just a phase shift, changing the "basic" period. A model equation could be

$$\ddot{x} + x - \varepsilon a x^3 = \varepsilon^2 \dot{x}(1 - x^2)$$

with a a suitable constant. We shall analyse this equation later on.

12.5.1 Procedure for Second-Order Calculation

Consider the n-dimensional equation in the standard form

$$\dot{x} = \varepsilon f(t, x) + \varepsilon^2 g(t, x) + \varepsilon^3 R(t, x, \varepsilon),$$

in which the vector fields f and g are T-periodic in t with averages f^0 and g^0 (T independent of ε). The vector fields f and g have to be sufficiently smooth; in particular, f has to be expanded, as we shall see. The higher-order term $R(t, x, \varepsilon)$ is smooth and bounded as ε tends to zero. If we are looking for T-periodic solutions, R in addition has to be T-periodic.

The interpretation of results of second-order averaging is more subtle than in first order, so we have to show some of the technical details of the construction. We denote with $\nabla f(t, x)$ the derivative with respect to x only; this is an $n \times n$ matrix. For instance, if $n = 2$, we have $x = (x_1, x_2), f = (f_1, f_2)$, and

$$\nabla f(t, x) = \begin{pmatrix} \frac{\partial f_1(t,x_1,x_2)}{\partial x_1} & \frac{\partial f_1(t,x_1,x_2)}{\partial x_2} \\ \frac{\partial f_2(t,x_1,x_2)}{\partial x_1} & \frac{\partial f_2(t,x_1,x_2)}{\partial x_2} \end{pmatrix}.$$

The reason we need this nabla operator (∇) is that to obtain a second-order approximation we have to Taylor expand the vector field. We need another vector field:

$$u^1(t, x) = \int_0^t (f(s, x) - f^0(x))ds - a(x),$$

where of course $f(s,x) - f^0(x)$ has average zero, but this does not hold necessarily for the integral (think of the function $\sin t \cos t$); $a(x)$ is chosen such that the average of u^1, $u^{10}(x)$ vanishes. We now introduce the *near-identity transformation*

$$x(t) = w(t) + \varepsilon u^1(t, w(t)). \tag{12.2}$$

This is also called the *averaging* or *normalising transformation.* Substituting Eq. (12.2) into the equation for x, we get

$$\dot{w}(t) + \varepsilon \frac{\partial u^1}{\partial t}(t, w(t)) + \varepsilon \nabla u^1(t, w(t))\dot{w}(t)$$

$$= \varepsilon f(t, w(t) + \varepsilon u^1(t, w(t))) + \varepsilon^2 g(t, w(t) + \varepsilon u^1(t, w(t))) + \varepsilon^3 \cdots .$$

Using the definition of u^1, the left-hand side of this equation becomes

$$(I + \varepsilon \nabla u^1(t, w))\dot{w} + \varepsilon f(t, w) - \varepsilon f^0(w),$$

where I is the identity $n \times n$ matrix. Inverting the matrix $(I + \varepsilon \nabla u^1(t, w))$ and expanding f and g, we obtain

$$\dot{w} = \varepsilon f^0(w) + \varepsilon^2 \nabla f(t, w) u^1(t, w) + \varepsilon^2 g(t, w) + \varepsilon^3 \cdots .$$

We put

$$f_1(t, x) = \nabla f(t, x) u^1(t, x),$$

the product of a matrix and a vector. Introducing now also the average f_1^0 and the equation

$$\dot{v} = \varepsilon f^0(v) + \varepsilon^2 f_1^0(v) + \varepsilon^2 g^0(v), \;\; v(0) = x(0),$$

we can prove that

$$x(t) = v(t) + \varepsilon u^1(t, v(t)) + O(\varepsilon^2)$$

on the timescale $1/\varepsilon$.

Note, that we did not simply expand with the first-order approximation as a first term; $v(t)$ contains already terms $O(\varepsilon)$ and $O(\varepsilon^2)$. What does this mean for the timescales? One would expect εt and $\varepsilon^2 t$, but this conclusion is too crude, as we shall see later on.

This second-order calculations has another interesting aspect. As u^1 is uniformly bounded, $v(t)$ is an $O(\varepsilon)$-approximation of $x(t)$. In general, it is different from the first-order approximation that we obtained before. This illustrates the nonuniqueness of asymptotic approximations. In some cases, however, the interpretation will be that $v(t)$ is an $O(\varepsilon)$-approximation on a different timescale. We shall return to this important point in a later section.

As mentioned above, we are now able to compute more accurate approximations but, more excitingly, we are able to discover new qualitative phenomena. A simple example is given below.

Example 12.18
Consider the equation

$$\ddot{x} + x - \varepsilon a x^3 = \varepsilon^2 \dot{x}(1 - x^2)$$

with a a suitable constant independent of ε. Introducing amplitude-phase variables $x(t) = r(t)\cos(t + \psi(t)), \dot{x}(t) = -r(t)\sin(t + \psi(t))$, we find

$$f(t, r, \psi) = \begin{pmatrix} -ar^3 \sin(t+\psi)\cos^3(t+\psi) \\ -ar^2 \cos^4(t+\psi) \end{pmatrix}, \ f^0(r, \psi) = \begin{pmatrix} 0 \\ -\frac{3}{8}ar^2 \end{pmatrix}.$$

Using initial values $r(0) = r_0, \psi(0) = 0$, we have a first-order approximation $x_a(t) = r_0 \cos(t - \varepsilon\frac{3}{8}ar_0^2 t), \dot{x}_a(t) = -r_0 \sin(t - \varepsilon\frac{3}{8}ar_0^2 t)$ corresponding with a set of periodic solutions with an $O(\varepsilon)$ shifted period that also depends on the initial r_0. Does a periodic solution branch off from one of the first-order approximations when considering second-order approximations? We have to compute and average $f_1 = \nabla f u^1$. Abbreviating $t + \psi = \alpha$, we find

$$\nabla f(t, r, \psi) = \begin{pmatrix} -3ar^2 \sin\alpha \cos^3\alpha & -ar^3\cos^4\alpha + 3ar^3 \sin^2\alpha\cos^2\alpha \\ -2ar\cos^4\alpha & 4ar^2\cos^3\alpha\sin\alpha \end{pmatrix},$$

and, using $\cos^4\alpha = \frac{3}{8} + \frac{1}{2}\cos 2\alpha + \frac{1}{8}\cos 4\alpha$,

$$u^1(t, r, \psi) = \begin{pmatrix} \frac{1}{8}ar^3\cos 2\alpha + \frac{1}{32}ar^3\cos 4\alpha \\ -\frac{1}{4}ar^2 \sin 2\alpha - \frac{1}{32}ar^2\sin 4\alpha \end{pmatrix}.$$

Multiplying, we find f_1, and after averaging

$$f_1^0(r, \psi) = \begin{pmatrix} 0 \\ -\frac{51}{256}a^2 r^4 \end{pmatrix}.$$

Now we can write down the equation for v as formulated in the procedure for second-order calculation. Note that $g^0(v)$ was calculated in Example 11.3 and we obtain

$$\dot{v}_1 = \varepsilon^2 \frac{v_1}{2}\left(1 - \frac{1}{4}v_1^2\right)$$
$$\dot{v}_2 = -\varepsilon\frac{3}{8}av_1^2 - \varepsilon^2\frac{51}{256}a^2v_1^4.$$

If $v_1(0) = r_0 = 2$, we have a stationary solution $v_1(t) = 2$, and with $v_2(0) = \psi(0) = 0$

$$v_2(t) = -\varepsilon\frac{3}{2}at - \varepsilon^2\frac{51}{16}a^2 t.$$

An $O(\varepsilon^2)$-approximation of the periodic solution is obtained by inserting this into u^1 so that

$$r(t) = 2 + \varepsilon a \cos 2(t + v_2(t)) + \varepsilon\frac{a}{4}\cos 4(t + v_2(t)) + O(\varepsilon^2),$$
$$\psi(t) = v_2(t) - \varepsilon a \sin 2(t + v_2(t)) - \varepsilon\frac{a}{8}\sin 4(t + v_2(t)) + O(\varepsilon^2),$$

valid on the timescale $1/\varepsilon$. At this order of approximation, the timescales for the periodic solution are $t, \varepsilon t$, and $\varepsilon^2 t$. Can you find the timescales for the other solutions?

12.5.2 An Unexpected Timescale at Second-Order

As another example, we consider the following Mathieu equation.

Example 12.19
Consider

$$\ddot{x} + (1 + \varepsilon a + \varepsilon^2 b + \varepsilon \cos 2t)x = 0$$

with free parameters a, b, which is a slight variation of the Mathieu equation we studied in Examples 10.9 and 11.6; see also the discussion in Section 15.3. Transforming by Eqs. 11.9

$$x(t) = y_1(t)\cos t + y_2(t)\sin t,\ \dot{x}(t) = -y_1(t)\sin t + y_2(t)\cos t,$$

we obtain the slowly varying system

$$\begin{aligned}
\dot{y}_1 &= \sin t(\varepsilon a + \varepsilon \cos 2t + \varepsilon^2 b)(y_1(t)\cos t + y_2(t)\sin t),\\
\dot{y}_2 &= -\cos t(\varepsilon a + \varepsilon \cos 2t + \varepsilon^2 b)(y_1(t)\cos t + y_2(t)\sin t).
\end{aligned}$$

We use again the terminology of the procedure for second-order calculation. We have to first order

$$f^0(y_1, y_2) = \begin{pmatrix} \frac{1}{2}(a - \frac{1}{2})y_2 \\ -\frac{1}{2}(a + \frac{1}{2})y_1 \end{pmatrix}$$

so that the first-order approximation is described by

$$\begin{aligned}
\dot{y}_{1a} &= \varepsilon \frac{1}{2}\left(a - \frac{1}{2}\right) y_{2a},\\
\dot{y}_{2a} &= -\varepsilon \frac{1}{2}\left(a + \frac{1}{2}\right) y_{1a}.
\end{aligned}$$

This is a system of linear equations with constant coefficients; the solutions are of the form $c\exp(\lambda t)$ with λ an eigenvalue of the matrix of coefficients. We find

$$\lambda_{1,2} = \pm \frac{1}{2}\sqrt{\frac{1}{4} - a^2}.$$

It is clear that we have stability of the trivial solution if $a^2 > \frac{1}{4}$ and instability if $a^2 < \frac{1}{4}$; $a = \pm\frac{1}{2}$ determines the boundary of the instability domain, which is called a Floquet tongue. On this boundary, the solutions are periodic to first approximation. What happens when we look more closely at the Floquet tongue? We find for ∇f and u^1

$$\nabla f(t, y_1, y_2) = \begin{pmatrix} \sin t \cos t(a + \cos 2t) & \sin^2 t(a + \cos 2t) \\ -\cos^2 t(a + \cos 2t) & -\sin t \cos t(a + \cos 2t) \end{pmatrix},$$

$$u^1(t, y_1, y_2) = \begin{pmatrix} -y_1(\frac{a}{4}\cos 2t + \frac{1}{16}\cos 4t) + y_2(-\frac{a}{4}\sin 2t + \frac{1}{4}\sin 2t - \frac{1}{16}\sin 4t) \\ -y_1(\frac{a}{4}\sin 2t + \frac{1}{4}\sin 2t + \frac{1}{16}\sin 4t) + y_2(\frac{a}{4}\cos 2t + \frac{1}{16}\cos 4t) \end{pmatrix}.$$

After some calculations, we find the average

$$f_1^0(y_1, y_2) = \begin{pmatrix} (\frac{a}{8} - \frac{a^2}{8} - \frac{1}{64})y_2 \\ (\frac{a}{8} + \frac{a^2}{8} + \frac{1}{64})y_1 \end{pmatrix}.$$

When adding g^0, we can write down the equation for v as formulated in the procedure for second-order calculation:

$$\dot{v}_1 = \varepsilon\frac{1}{2}\left(a - \frac{1}{2}\right)v_2 + \varepsilon^2\left(\frac{a}{8} - \frac{a^2}{8} - \frac{1}{64} + \frac{1}{2}b\right)v_2,$$

$$\dot{v}_2 = -\varepsilon\frac{1}{2}\left(a + \frac{1}{2}\right)v_1 + \varepsilon^2\left(\frac{a}{8} + \frac{a^2}{8} + \frac{1}{64} - \frac{1}{2}b\right)v_1.$$

Choosing for instance $a = \frac{1}{2}$, we have one of the boundaries of the Floquet tongue consisting of periodic solutions. The equations for v become

$$\dot{v}_1 = \varepsilon^2\left(\frac{1}{64} + \frac{1}{2}b\right)v_2,$$

$$\dot{v}_2 = -\varepsilon\frac{1}{2}v_1 + \varepsilon^2\left(\frac{7}{64} - \frac{1}{2}b\right)v_1.$$

For the eigenvalues of the matrix of coefficients to second order, we find

$$\lambda_{1,2} = \pm\sqrt{-\frac{1}{4}\left(\frac{1}{32} + b\right)\varepsilon^3 + \left(\frac{1}{64} + \frac{1}{2}b\right)\left(\frac{7}{64} - \frac{1}{2}b\right)\varepsilon^4}.$$

We conclude that if $\frac{1}{32} + b > 0$, we have stability, and if $\frac{1}{32} + b < 0$, we have instability. The value $b = -\frac{1}{32}$ gives us the second-order approximation of the Floquet tongue. This result has an interesting consequence for the timescales of the solutions in the case $a = \frac{1}{2}$; they are $t, \varepsilon t, \varepsilon^{\frac{3}{2}}t, \varepsilon^2 t$. The timescale $\varepsilon^{\frac{3}{2}}t$ is quite unexpected.

Remark

Second-order general averaging, without periodicity assumptions, runs along the same lines as demonstrated in the periodic case. For the theory and examples, see Sanders and Verhulst (1985).

12.6 Approximations Valid on Longer Timescales

In a number of problems, we have a priori knowledge that the solutions of equations we are studying exist on a longer timescale than $1/\varepsilon$ or even exist for all time. Is it not possible in these cases to obtain approximations valid on such a longer timescale? For instance, when calculating an $O(\varepsilon^2)$-approximation on the timescale $1/\varepsilon$ as in the preceding section, can we not as a trade-off consider this as an $O(\varepsilon)$-approximation on the timescale $1/\varepsilon^2$? It turns out that in general the answer is "no," as can easily be seen from examples. However, we shall consider an important case where this idea carries through.

12.6.1 Approximations Valid on $O(1/\varepsilon^2)$

We formulate the following result.

Theorem 12.2
Consider again the n-dimensional equation in the standard form

$$\dot{x} = \varepsilon f(t, x) + \varepsilon^2 g(t, x) + \varepsilon^3 R(t, x, \varepsilon)$$

in which the vector fields f and g are T-periodic in t with averages f^0 and g^0 (T independent of ε); f, g, and R are sufficiently smooth. Suppose that

$$f^0(x) = 0.$$

We have from the first-order approximation $x(t) = x(0) + O(\varepsilon)$ on the timescale $1/\varepsilon$. Following the construction of the second-order approximation of the preceding section, we have with $f^0(x) = 0$ the equation

$$\dot{v} = \varepsilon^2 g^0(v), \; v(0) = x(0).$$

It is easy to prove that

$$x(t) = v(t) + O(\varepsilon)$$

on the timescale $1/\varepsilon^2$ (Van der Burgh, 1975).

A simple example is given as follows.

Example 12.20

$$\ddot{x} + x - \varepsilon a x^2 = 0, \; r(0) = r_0, \psi(0) = 0,$$

with a an ε-independent parameter. In amplitude-phase variables r, ψ, we find with transformation (11.5)

$$\begin{aligned} \dot{r} &= -ar^2 \sin(t + \psi) \cos^2(t + \psi), \\ \dot{\psi} &= -ar \cos^3(t + \psi), \end{aligned}$$

so with $f(t, r, \psi)$ for the right-hand side we have $f^0(r, \psi) = 0$. For the second-order approximation, we abbreviate $t + \psi = \alpha$ and calculate

$$\nabla f(t, r, \psi) = \begin{pmatrix} -2ar \sin\alpha \cos^2\alpha & -ar^2 \cos^3\alpha + 2ar^2 \sin^2\alpha \cos\alpha \\ -a \cos^3\alpha & 3ar \cos^2\alpha \sin\alpha \end{pmatrix},$$

$$u^1(t,r,\psi) = \begin{pmatrix} \frac{1}{3}ar^2\cos^3\alpha \\ -ar\sin\alpha + \frac{1}{3}ar\sin^3\alpha \end{pmatrix}.$$

With $f_1 = \nabla f u^1$, we find after averaging

$$f_1^0(r,\psi) = \begin{pmatrix} 0 \\ -\frac{5}{12}a^2r^2 \end{pmatrix}.$$

We conclude that $\dot{v}_1 = 0, \dot{v}_2 = -\varepsilon^2\frac{5}{12}a^2v_1^2$, and we have $x(t) = r_0\cos(t - \varepsilon^2\frac{5}{12}a^2r_0^2t) + O(\varepsilon)$, valid on the timescale $1/\varepsilon^2$.

This is already a nontrivial result, but a more interesting example arises when considering the Mathieu equation for other resonance values than those studied in Example 12.19.

Example 12.21
Consider

$$\ddot{x} + (n^2 + \varepsilon a + \varepsilon^2 b + \varepsilon\cos mt)x = 0,$$

again with free parameters a, b. Transforming by Eq. (11.10)

$$x(t) = y_1(t)\cos nt + \frac{1}{n}y_2(t)\sin nt,\ \dot{x}(t) = -ny_1(t)\sin nt + y_2(t)\cos nt,$$

we obtain the slowly varying system

$$\dot{y}_1 = \frac{\varepsilon}{n}\sin nt(a+\cos mt)\left(y_1(t)\cos nt + \frac{1}{n}y_2(t)\sin nt) + O(\varepsilon^2\right),$$
$$\dot{y}_2 = -\frac{\varepsilon}{n}\cos nt(a+\cos mt)\left(y_1(t)\cos nt + \frac{1}{n}y_2(t)\sin nt\right) + O(\varepsilon^2).$$

It is easy to see (by averaging) that f^0 is nontrivial if $2n - m = 0$, which is the case of Example 12.19. This is the most prominent resonance of the Mathieu equation. For other rational ratios of m and n, we find other resonances with different sizes of the resonance tongues. As an example, we explore the case $m = n = 2$. If $2n - m \neq 0$, the averaged equations are dominated by the parameter a so it makes sense to choose $a = 0$, as the Floquet tongue will be narrower than in the case $2n - m = 0$. The equation becomes

$$\ddot{x} + (4 + \varepsilon^2 b + \varepsilon\cos 2t)x = 0.$$

We omit the expressions for ∇f and u^1 and produce f_1 directly:

$$f_1(t,y_1,y_2) = \begin{pmatrix} \frac{1}{192}\sin 4t(-6y_1 + 2y_1\cos 4t + y_2\sin 4t) \\ \\ \frac{-1}{48}\cos^2 2t(-6y_1 + 2y_1\cos 4t + y_2\sin 4t) \end{pmatrix}.$$

So we find

$$f_1^0(y_1, y_2) = \begin{pmatrix} \frac{1}{384} y_2 \\ \\ -\frac{1}{96} y_1 \end{pmatrix}$$

and for the second-order equations

$$\dot{v}_1 = \varepsilon^2 \frac{y_2}{8} \left(\frac{1}{48} + b \right), \ \dot{v}_2 = \varepsilon^2 \frac{y_1}{2} \left(\frac{5}{48} - b \right).$$

The solutions $v_1(t)$ and $v_2(t)$ are $O(\varepsilon)$-approximations of $y_1(t)$ and $y_2(t)$ valid on the timescale $1/\varepsilon^2$. From the eigenvalues, we conclude instability if

$$-\frac{1}{48} < b < \frac{5}{48}.$$

The Floquet tongue in the case $m = n = 2$ is determined by the boundary values $4 - \frac{1}{48}\varepsilon^2$ and $4 + \frac{5}{48}\varepsilon^2$.

When assuming $m \neq 2n$ and $m \neq n$, we can study higher order Floquet tongues. See also the discussion in Example 10.9, where the tongues are determined by the continuation (Poincaré-Lindstedt) method, and in particular Fig. 10.1.

12.6.2 Timescales near Attracting Solutions

Another natural idea to obtain extension of the timescale of validity is attraction. Suppose the solutions are attracted exponentially fast to a special solution, equilibrium or periodic, for which the process will also result in a kind of compression of the neighbouring solutions. In this case, we have the following result.

Theorem 12.3
Consider the n-dimensional equation in the standard form

$$\dot{x} = \varepsilon f(t, x) + \varepsilon^2 R(t, x, \varepsilon)$$

in which the vector field f is T-periodic in t with average f^0 (T independent of ε); f and R are sufficiently smooth. Suppose that $f^0(x)$ contains a critical point (equilibrium of the averaged equation) $x = a$, so $f^0(a) = 0$. We assume that all the eigenvalues in $x = a$ have a negative real part. The solution $x(t)$ starting in $x(0)$, which is located in an interior subset of the domain of attraction of $x = a$, is approximated by the solution $x_a(t)$ of the averaged equation starting in $x(0)$ as

$$x(t) - x_a(t) = O(\varepsilon), \ 0 \leq t < \infty.$$

Example 12.22
A simple example is the one-dimensional equation

$$\dot{x} = -\varepsilon 2 \sin^2 tx + \varepsilon^2 R(t, x)$$

with averaged equation

$$\dot{x}_a = -x_a.$$

We have $x(t) = x(0)\exp(-t) + O(\varepsilon)$ for all time.

Example 12.23
More important is the forced Duffing equation in Example 11.4

$$\ddot{x} + \varepsilon\mu\dot{x} + \varepsilon\gamma x^3 + x = \varepsilon h \cos\omega t,$$

with $\mu > 0$; we choose the case of exact resonance $\omega = 1$. In this case, the system averaged to first order is

$$\dot{r}_a = -\frac{1}{2}\varepsilon(\mu r_a + h \sin\psi_a),$$
$$\dot{\psi}_a = -\frac{1}{2}\varepsilon\left(-\frac{3}{4}\gamma r_a^2 + h\frac{\cos\psi_a}{r_a}\right).$$

The critical points are determined by the equations

$$h \sin\psi_a = -\mu r_a,\ \ h\cos\psi_a = \frac{3}{4}\gamma r_a^3.$$

Using these equations, it is not difficult to find the eigenvalues

$$\lambda_{1,2} = -\frac{1}{2}\mu \pm i\frac{3\sqrt{3}}{8}|\gamma| r_a^2,$$

so if we find critical points, they are asymptotically stable and the solutions attracting to the corresponding periodic solution are approximated by the solutions of the averaged equation for all time.

A problem arises when the original equation is autonomous. In this case, a periodic solution has at least one eigenvalue zero, a feature that is inherited by the averaged equation.

Example 12.24
Consider again the Van der Pol equation

$$\ddot{x} + x = \varepsilon(1 - x^2)\dot{x}$$

with averaged equation (amplitude-phase)

$$\dot{r}_a = \varepsilon\frac{r_a}{2}\left(1 - \frac{1}{4}r_a^2\right),\ \dot{\psi}_a = 0.$$

As we remarked in the discussion of Example 11.3, we can reduce such an autonomous equation by introducing $\tau = t + \psi$ and average over τ. We find

$$\frac{dr_a}{d\tau} = \varepsilon \frac{r_a}{2}\left(1 - \frac{1}{4} r_a^2\right),$$

and we can apply the theory as follows. Solutions $r(\tau)$ starting outside a neighbourhood of $r = 0$ are approximated for all times τ or t by the solutions of the averaged equation. Such a result is not valid for the phase ψ.

This example can easily be generalised for second-order autonomous equations.

12.7 Identifying Timescales

A fundamental question that comes up very often is whether we have the right expansion coefficients with respect to ε and, in the context of evolution problems, whether we have the right timescales. In the framework of boundary layer problems, we have developed a technique in Chapter 4 to identify local or boundary layer variables. This technique works fine in many cases but not always. When studying evolution problems, the situation is worse; we shall first review some clarifying examples.

12.7.1 Expected and Unexpected Timescales

In Chapters 10 and 11, we have seen that in equations of the form $\dot{x} = f(t, x, \varepsilon)$ where the right-hand side depends smoothly on ε, the solutions can sometimes be expanded in powers of ε while we have timescales such as $t, \varepsilon t$, and $\varepsilon^2 t$. In fact, as we have shown before, when first-order averaging is possible, we have definitely the timescales t and εt. However, in general and certainly at second order, the situation is not always as simple as that.

Example 12.25
Consider for $t \geq 0$ the linear initial value problem

$$\ddot{x} + \frac{1}{1+t}\dot{x} = \varepsilon \frac{2}{(1+t)^2},\ x(0) = \dot{x}(0) = 0,$$

with solution

$$x(t) = \varepsilon \ln^2(1+t),$$

so $x(t)$ is characterised by the timescale $\varepsilon \ln^2(1+t)$, and $\dot{x}(t)$ by the timescales t and $\varepsilon \ln(1+t)$.

Example 12.26
Consider the initial value problem for the Bernoulli equation

$$\dot{x} = x^\alpha - cx,\ x(0) = 1.$$

For the constants, we have $0 < \alpha < 1, c > 0$. The solution of the initial value problem is

$$x(t) = \left(\frac{1}{c} + \left(1 - \frac{1}{c}\right) e^{-c(1-\alpha)t}\right)^{\frac{1}{1-\alpha}}.$$

The solutions starting with $x(0) > 0$ tend to stable equilibrium $c^{-\frac{1}{1-\alpha}}$. We consider two cases.

1. Choose $c = \varepsilon$, where α does not depend on ε:

$$\dot{x} = x^\alpha - \varepsilon x, \ x(0) = 1,$$

   with solution

$$x(t) = \left(\frac{1}{\varepsilon} + \left(1 - \frac{1}{\varepsilon}\right) e^{-\varepsilon(1-\alpha)t}\right)^{\frac{1}{1-\alpha}}.$$

   One of the expansion coefficients is $\varepsilon^{-\frac{1}{1-\alpha}}$, and the relevant timescale is εt.
2. Choose $c = \varepsilon, 1 - \alpha = \varepsilon$, producing

$$\dot{x} = x^{1-\varepsilon} - \varepsilon x, \ x(0) = 1,$$

   with solution

$$x(t) = \left(\frac{1}{\varepsilon} + \left(1 - \frac{1}{\varepsilon}\right) e^{-\varepsilon^2 t}\right)^{\frac{1}{\varepsilon}}.$$

   The relevant timescale is $\varepsilon^2 t$, and $1/\varepsilon$ plays a part in the expansion.

The Mathieu equation in Example 12.19 is important to show that unexpected timescales occur in practical problems.We summarise as follows.

Example 12.27
Consider

$$\ddot{x} + (1 + \varepsilon a + \varepsilon^2 b + \varepsilon \cos 2t)x = 0.$$

On choosing $a = \pm\frac{1}{2}$, the solutions are located near or on the Floquet tongue. A second-order approximation has shown that the timescales in this case are $t, \varepsilon t, \varepsilon^{\frac{3}{2}t}, \varepsilon^2 t$. They naturally emerge from the averaging process.

Other important examples of unexpected timescales can be found in the theory of Hamiltonian systems. These examples are much more complicated.

12.7.2 Normal Forms, Averaging and Multiple Timescales

The Mathieu equation that we discussed above is a linear equation where the calculation to second order yields an unexpected timescale. In Section 15.3, we show for matrices A with constant entries that, when arising in equations of the form $\dot{x} = Ax$, in particular the bifurcation values produce such unexpected timescales. For nonlinear equations, we have no general theory, only examples and local results.

Where does this leave us in constructing approximations for initial value problems? When given a perturbation problem, we always start with transformations and other operations that we think suitable for the problem. The main directive is that there should be *no a priori assumptions on the timescales.* The general framework for this is the theory of normal forms, of which averaging is one of the parts (see for instance Sanders and Verhulst, 1985). In this framework, suitable transformations are introduced, and from the analysis of the resulting equations the timescales follow naturally. Also, the corresponding proofs of validity yield confirmation and natural restrictions on the use of these timescales.

It should be clear by now that the method of multiple timescales for initial value problems of ordinary differential equations is not suitable for research problems except in the case of simple problems that are accessible also to first-order averaging. For other problems, this method presupposes the presence of certain timescales, a knowledge we simply do not have.

Multiple timing, as it is often called, is of course suitable and an elegant method in the cases of solved problems where we know a priori the structure of the approximations. It should also be mentioned that the formal calculation schemes of multiple timing are easier to extend to problems for partial differential equations; we will return to this in Chapter 14. However, proofs of validity of such extensions are often lacking.

12.8 Guide to the Literature

There are thousands of papers associated with this chapter. We will mention here basic literature with good literature sections for further study. Several times we touched upon the relation between averaging and normalisation. Averaging can be seen as a special normal-form method with the advantage of explicitly formulated normal forms. On the other hand, normal-form methods are approximation tools, using in some form localisation around a special point or another solution. More about this relation can be found in Arnold (1982) and Sanders and Verhulst (1985). The same references can be used for the theory of averaging over angles. A monograph by Lochak and Meunier (1988) is devoted to this topic.

Invariant manifolds represent a subject that is still very much in development. The idea to obtain tori by continuation was introduced by Bogoliubov and Mitropolsky (1961, 1963) and extended by Hale (1969). Tori can emerge in a different way by Hopf bifurcation of a periodic solution; this is called Neimark-Sacker bifurcation. Usually, numerical bifurcation path-following programs are used to pinpoint such bifurcations, but the demonstration by averaging is for instance shown in Bakri et al. (2004).

Invariant tori in the Hamiltonian context is a very large subject with an extensive body of literature and many interesting books; for an introduction, see Arnold (1978) or Verhulst (2000). Lochak et al. (2003) discusses what happens between the tori and gives many references. A survey of normalisation and averaging for Hamiltonian systems is given in Verhulst (1998).

The theory of adiabatic invariants is a classical subject that is tied in both with problems of averaging over angles and bifurcation theory. For the classical theory, see Arnold (1978) and, in particular, the survey by Henrard (1993). Slowly varying coefficients may lead a system to passage through a bifurcation. Neishstadt (1986, 1991), Cary et al. (1986), Haberman (1978) and Bourland and Haberman (1990) analysed the slow passage through a separatrix. See also Diminnie and Haberman (2002) for changes of the adiabatic invariant in such a setting. Adiabatic changes in a Hamiltonian system with two degrees of freedom are discussed in Verhulst and Huveneers (1998).

Second-order averaging and longer timescales are studied in Sanders and Verhulst (1985). Van der Burgh (1975) produced the first estimates on the timescale $1/\varepsilon^2$. An extension to $O(\varepsilon^2)$ on the timescale $1/\varepsilon^2$ is given in Verhulst (1988).

Multiple-timescale methods are discussed for instance in Hinch (1991), Kevorkian and Cole (1996), and Holmes (1998). For a comparison of multiple timing and averaging, see Perko (1969) and Kevorkian (1987).

12.9 Exercises

Exercise 12.1

$$\ddot{x} + 2\varepsilon\dot{x} + \varepsilon\dot{x}^3 + x = 0.$$

Use the amplitude-angle transformation 12.1 to obtain a system that can be averaged over the angle ϕ; give the result for the approximation of the amplitude.

Exercise 12.2

$$\ddot{x} + \varepsilon\mu\dot{x} + \omega^2(\varepsilon t)x + \varepsilon x^3 = 0.$$

Use the amplitude-angle transformation (12.1) to obtain a system that can be averaged over the angle ϕ; give the result (with additional assumptions) for the approximation of the amplitude.

Exercise 12.3 Consider the system

$$\begin{aligned}
\dot{x} &= \varepsilon + \varepsilon\sin(\phi_1 - \phi_2),\\
\dot{\phi}_1 &= x,\\
\dot{\phi}_2 &= x^2.
\end{aligned}$$

Determine the location(s) of the resonance manifold(s) and an approximation away from these location(s).

Exercise 12.4

$$\begin{aligned}\dot{x} &= \varepsilon\cos(t - \phi_1 + \phi_2) + \varepsilon\sin(t + \phi_1 + \phi_2),\\ \dot{\phi}_1 &= 2x,\\ \dot{\phi}_2 &= x^2.\end{aligned}$$

Determine the location(s) of the resonance manifold(s) and an approximation away from these location(s).

Exercise 12.5 In Example 12.9, we stated that in the two resonance zones the flow is described by two pendulum equations. Verify this statement.

Exercise 12.6 Consider the Duffing equation with slowly varying coefficients in the form

$$\ddot{x} + \omega^2(\varepsilon t)x = \varepsilon a(\varepsilon t)x^3.$$

Compute an adiabatic invariant for the equation with suitable assumptions on the coefficients $a(\varepsilon t)$ and $\omega(\varepsilon t)$.

Exercise 12.7 In the beginning of Section 12.5, we stated that the equation

$$\ddot{x} + \varepsilon\mu\dot{x} + x = \varepsilon^2\dot{x}(1 - x^2)$$

with positive damping constant μ does not contain a periodic solution. Ignoring the theory of periodic solutions behind this, verify this statement by discussing the second-order approximation.

Exercise 12.8 Determine a second-order approximation for the solutions of

$$\ddot{x} + x - \varepsilon a x^2 = \varepsilon^2\dot{x}(1 - x^2)$$

with a a suitable constant independent of ε. Can we identify a periodic solution? Replace the $O(\varepsilon)$ term by εx^m with m an even number and repeat the analysis.

Exercise 12.9 A start for the calculation of higher-order Floquet tongues in the case of the Mathieu equation from Example 12.19 with $m \neq 2n$ and $m \neq n$ is the determination of the second-order approximation. Show that the equations are

$$\dot{v}_1 = \varepsilon^2\frac{v_2}{2n^2}\left(\frac{1}{2m^2 - 8n^2} + b\right),\ \dot{v}_2 = -\varepsilon^2\frac{v_1}{2}\left(\frac{1}{2m^2 - 8n^2} + b\right).$$

What can we conclude at this stage for the size of the higher-order Floquet tongues? Note in this context that the first tongue has a separation between the boundaries that are to first order straight lines, the second tongue is bounded by parabolas, and the higher-order tongues have boundaries that are tangent as ε tends to zero; see Fig. 10.1.

Exercise 12.10 Consider the system

$$\dot{x} = y + \varepsilon(x^2 \sin 2t - \sin 2t),$$
$$\dot{y} = -4x.$$

Find equilibria and corresponding periodic solutions of the associated averaged system. For which initial conditions can we extend the timescale of validity beyond $1/\varepsilon$?

Exercise 12.11 Consider a Hamiltonian system with two degrees of freedom, the so-called Hénon-Heiles family:

$$\ddot{x} + x = \varepsilon(a_1 x^2 + a_2 y^2),$$
$$\ddot{y} + \omega^2 y = \varepsilon 2 a_2 xy,\ a_2 \neq 0.$$

a. Show that for $\omega = 2$ first-order averaging produces a nontrivial result.
b. Consider $\omega = 1$ and determine the equations for the second-order approximation.
c. Determine in the case $\omega = 1$ two integrals of motion of the averaged equations and indicate their geometrical meaning.
d. Determine in the case $\omega = 1$ the conditions for the existence of short-periodic solutions in a general position away from the normal modes.

Exercise 12.12 Consider a Hamiltonian system with two degrees of freedom to which we have added damping and parametric excitation:

$$\ddot{x} + 2\varepsilon\dot{x} + (4 + \varepsilon a \cos t)x = \varepsilon y^2,$$
$$\ddot{y} + 2\varepsilon\dot{y} + (1 + \varepsilon b \cos 2t)y = 2\varepsilon xy.$$

a. If $\varepsilon = 0$, we have two normal modes. Can we continue them for $\varepsilon > 0$?
b. Can you find other periodic solutions?