9. Time-Independent Canonical Perturbation Theory

First we consider the perturbation calculation only to first order, limiting ourselves to only one degree of freedom. Furthermore, the system is to be conservative, $\partial H/\partial t = 0$, and periodic in both the unperturbed and perturbed case. In addition to periodicity, we shall require the Hamilton-Jacobi equation to be separable for the unperturbed situation. The unperturbed problem $H_0(J_0)$ which is described by the action-angle variables J_0 and w_0 will be assumed to be solved. Thus we have, for the unperturbed frequency:

$$\nu_0 = \frac{\partial H_0}{\partial J_0} \tag{9.1}$$

and

$$w_0 = \nu_0 t + \beta_0 \ . \tag{9.2}$$

Then the new Hamiltonian reads, up to a perturbation term of first order:

$$H = H_0\big(J_0\big) + \varepsilon H_1\big(w_0, J_0\big) \ , \tag{9.3}$$

where ε is a small parameter. Our goal now is to find a canonical transformation from the action-angle variables (J_0, w_0) of the unperturbed problem $H_0(J_0)$ to action-angle variables (J, w) of the total problem $H(J)$; this canonical transformation should make the perturbed problem become solvable. If we can achieve this, then it holds that $H = E(J)$, where $J = \text{const.}$, and now (1) and (2) are replaced by

$$\nu = \frac{\partial H(J)}{\partial J} \tag{9.4}$$

and

$$w = \nu t + \beta \ . \tag{9.5}$$

The canonical transformation in question can be generated with the help of the generating function of the type $F_2(q, P) : W = W(w_0, J)$. w_0 stands for the old coordinate and J for the new momentum. Since we are limiting ourselves to $\partial F_2/\partial t = 0$, it holds that $H_{\text{old}} = H_{\text{new}}$. Then the Hamilton-Jacobi equation reads:

$$H\big(w_0, J_0\big) = H\left(w_0, \frac{\partial W}{\partial w_0}\right) = E(J) \tag{9.6}$$

with

$$J_0 = \frac{\partial W(w_0, J)}{\partial w_0}, \qquad w = \frac{\partial W(w_0, J)}{\partial J}. \tag{9.7}$$

This corresponds to the familiar transformation equations $p = \partial F_2/\partial q$ and $Q = \partial F_2/\partial P$.

It is important to emphasize that for the perturbed problem, the (w, J) are "good" action-angle variables, while the (w_0, J_0) "basis" no longer plays the role of action-angle variables. w_0 is angle variable for the unperturbed case and is related to the original coordinate q by

$$q = \sum_{k=-\infty}^{+\infty} a_k(J_0)\, \mathrm{e}^{2\pi i k w_0} \quad \text{(libration)}, \tag{9.8}$$

or

$$q - q_0 w_0 = \sum_{k=-\infty}^{+\infty} a_k(J_0)\, \mathrm{e}^{2\pi i k w_0} \quad \text{(rotation)}. \tag{9.9}$$

Certainly (w_0, J_0) remain canonical variables for the perturbed situation, since they are, according to the above, related to the original canonical variables (q, p) by a canonical transformation. J_0 is now, however, no longer constant $[\dot{J}_0 = -\partial H/\partial w_0 = -\varepsilon(\partial H_1(w_0, J_0)/\partial w_0)]$ and w_0 is no longer a linear function in time $[\dot{w}_0 = \partial H/\partial J_0 = \partial H_0(J_0)/\partial J_0 + \varepsilon(\partial H_1(w_0, J_0)/\partial J_0) \neq \text{const.}]$. Since (w, J) are action-angle variables, w increases by one unit when q runs through one period. This also applies, however, to w_0, because q is, according to (9.8), a periodic function of w_0 with period 1. The canonical transformation (9.8) expresses q in dependence of (w_0, J_0), and has nothing to do with the particular form of the Hamiltonian.

We now return to (9.6) and treat this equation perturbatively, i.e., we expand both sides:

$$H(w_0, J_0) = H_0(J_0) + \varepsilon H_1(w_0, J_0) + \ldots \tag{9.10}$$

$$E(J; \varepsilon) = E_0(J) + \varepsilon E_1(J) + \varepsilon^2 E_2(J) + \ldots. \tag{9.11}$$

We apply the same procedure to the generating function $W(w_0, J)$ of the canonical transformation (9.7), which transforms (w_0, J_0) to (w, J):

$$W(w_0, J) = \underbrace{W_0(w_0, J)}_{= w_0 J} + \varepsilon W_1(w_0, J) + \varepsilon^2 W_2(w_0, J) + \ldots. \tag{9.12}$$

For $\varepsilon = 0$, only the identity transformation $w_0 J$ remains. The transformation equations (9.7) take on the following form:

$$\begin{aligned} J_0 &= \frac{\partial W(w_0, J)}{\partial w_0} = J + \varepsilon \frac{\partial W_1(w_0, J)}{\partial w_0} + \ldots \\ w &= \frac{\partial W(w_0, J)}{\partial J} = w_0 + \varepsilon \frac{\partial W_1(w_0, J)}{\partial J} + \ldots. \end{aligned} \tag{9.13}$$

The Hamilton-Jacobi equation (9.6) can then be written in first-order perturbation theory:

$$H\left(w_0, \frac{\partial W}{\partial w_0}\right) = E(J):$$

$$H_0(J_0) + \varepsilon H_1\Bigg(w_0, \underbrace{\frac{\partial W}{\partial w_0}}_{(13): J+\varepsilon \partial W_1/\partial w_0}\Bigg) = E_0(J) + \varepsilon E_1(J) \tag{9.14}$$

and with

$$\begin{aligned} H_0(J_0) &\underset{(13)}{=} H_0\left(J + \varepsilon\frac{\partial W_1}{\partial w_0}\right) \\ &= H_0(J) + \varepsilon\frac{\partial W_1}{\partial w_0}\frac{\partial H_0(J)}{\partial J} + \mathrm{O}(\varepsilon^2) \end{aligned}$$

we get

$$\begin{aligned} &H_0(J) + \varepsilon\left[H_1\left(w_0, J\right) + \frac{\partial H_0(J)}{\partial J}\frac{\partial W_1(w_0, J)}{\partial w_0}\right] \\ &\quad = E_0(J) + \varepsilon E_1(J)\,. \end{aligned} \tag{9.15}$$

Here, only w_0 and the constant J still appear. Comparison of coefficients in ε finally yields:

$$\varepsilon^0: \; H_0(J) = E_0(J)\,, \tag{9.16}$$

$$\varepsilon^1: \; H_1\left(w_0, J\right) + \frac{\partial H_0(J)}{\partial J}\frac{\partial W_1(w_0, J)}{\partial w_0} = E_1(J)\,. \tag{9.17}$$

Equation (9.17) contains the two unknown functions $W_1(w_0, J)$ and $E_1(J)$. Two assumptions permit us to solve (9.17). First of all, we set

$$\frac{\partial H_0(J)}{\partial J} = \left.\frac{\partial H_0(J_0)}{\partial J_0}\right|_{J_0=J} = \nu_0\,. \tag{9.18}$$

ν_0 is the frequency of the solved problem! Then (9.17) becomes

$$H_1\left(w_0, J\right) + \nu_0\frac{\partial W_1(w_0, J)}{\partial w_0} = E_1(J)\,. \tag{9.19}$$

The inhomogeneous term H_1 is given, $E_1(J)$ is unknown. Thus, (9.19) is a linear partial differential equation with constant coefficient (ν_0) for W_1.

Next we take advantage of the fact that the function W_1 is a periodic function of w_0. In this respect we recall that the function $W^*(w_0, w) \equiv W(w_0, J) - w_0 J$ is a periodic function of w_0; since J is an action variable here, it holds that $J = \oint p\,dq = \oint(\partial W/\partial q)dq$, so that for a single rotation in q, the action increases by J. Simultaneously, w_0 increases by one unit, (9.8). Then it holds that

$$W^*(w_0+1, w) = W(w_0+1, J) - (w_0+1)J = W(w_0, J) + J - w_0 J - J$$
$$= W(w_0, J) - w_0 J = W^*(w_0, w) \,. \tag{9.20}$$

Because

$$W^* \underset{(12)}{=} \varepsilon W_1(w_0, J) + \varepsilon^2 W_2(w_0, J) + \dots \tag{9.21}$$

every W_i, in particular, W_1, is also a periodic function in w_0:

$$W_1(w_0, J) = \sum_{k=-\infty}^{+\infty} C_k(J)\, \mathrm{e}^{2\pi i k w_0} \,. \tag{9.22}$$

Consequently, $\partial W_1/\partial w_0$ in (9.19) contains no constant term. If one now averages (9.19) over one period w_0 of the unperturbed problem, one gets

$$E_1(J) = \overline{H_1(w_0, J)} \,, \tag{9.23}$$

because the average over the derivative of the periodic function (9.22) vanishes. If we then insert the expression for $E_1(J)$ in (9.23) into (9.19), we have

$$\nu_0 \frac{\partial W_1(w_0, J)}{\partial w_0} = -\left[H_1(w_0, J) - \overline{H_1(w_0, J)}\right] =: -\{H_1\} \,. \tag{9.24}$$

Here, the right-hand side is known, and we thus get a linear partial differential equation with constant coefficients for W_1. Note that averaging the right-hand side of (9.24) indeed yields zero.

If we now are interested in the new frequency, the knowledge of W_1 is superfluous, since we only need (9.23) in

$$\nu = \frac{\partial E(J)}{\partial J} = \nu_0 + \varepsilon \frac{\partial E_1(J)}{\partial J} = \nu_0 + \varepsilon \frac{\partial \bar{H}_1}{\partial J} \,. \tag{9.25}$$

We now come to a few simple illustrative examples and begin by determining the dependence of the frequency on the amplitude in first-order perturbation theory for a perturbed oscillator potential,

$$V(q) = \frac{k}{2} q^2 + \frac{1}{6} \varepsilon m q^6 \,, \tag{9.26}$$

where $k = m\omega_0^2$, and ω_0 is the small-amplitude frequency of the unperturbed oscillator. The Hamiltonian of the problem is given by

$$H = T + V = \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2 + \varepsilon \frac{m q^6}{6} = H_0 + \varepsilon H_1 \,. \tag{9.27}$$

For the unperturbed Hamiltonian we already have found that

$$H_0 = \nu_0 J_0 = \frac{\omega_0}{2\pi} J_0 \,, \qquad w_0 = \nu_0 t + \beta_0 \,, \tag{9.28}$$

$$q = \sqrt{\frac{J_0}{\pi m \omega_0}} \sin(2\pi w_0) \,, \qquad p = \sqrt{\frac{m\omega_0 J_0}{\pi}} \cos(2\pi w_0) \,. \tag{9.29}$$

According to (9.23), we have to compute

$$E_1(J) = \bar{H}_1 = \frac{m}{6}\overline{q^6} = \frac{m}{6}\left(\frac{J}{\pi m \omega_0}\right)^3 \overline{\sin^6(2\pi w_0)} \,. \tag{9.30}$$

In order to determine the average value in (9.30), we recall that

$$\begin{aligned}
\sin^6\alpha &= \left(\frac{1}{2\mathrm{i}}\right)^6 \left(\mathrm{e}^{\mathrm{i}\alpha} - \mathrm{e}^{-\mathrm{i}\alpha}\right)^6 = \left(-\frac{1}{4}\right)^3 \\
&\quad \times \left[1 \cdot \mathrm{e}^{6\mathrm{i}\alpha} - 6\mathrm{e}^{4\mathrm{i}\alpha} + 15\mathrm{e}^{2\mathrm{i}\alpha} - 20 + 15\mathrm{e}^{-2\mathrm{i}\alpha} + 6\mathrm{e}^{-4\mathrm{i}\alpha} + 1 \cdot \mathrm{e}^{-6\mathrm{i}\alpha}\right] \\
&= -\frac{2}{64}[\cos(6\alpha) - 6\cos(4\alpha) + 15\cos(2\alpha) - 10] \,.
\end{aligned}$$

Thus we have

$$\overline{\sin^6\alpha} = \left(-\tfrac{2}{64}\right)(-10) = \tfrac{5}{16} \,. \tag{9.31}$$

For the energy correction E_1 in $E = E_0 + \varepsilon E_1$, it therefore follows from (9.30) that

$$E_1(J) = \frac{m}{6}\,\frac{5}{16}\left(\frac{J}{\pi m \omega_0}\right)^3 \,. \tag{9.32}$$

We have been looking for the new frequency,

$$\nu = \frac{\partial E(J)}{\partial J} = \nu_0 + \varepsilon\frac{5m}{32}\,\frac{J^2}{(\pi m \omega_0)^3} \,. \tag{9.33}$$

If A is the maximum amplitude of the unperturbed harmonic oscillator, then $J = J_0 = \pi m \omega_0 A^2$ in first-order perturbation theory. Then (9.33) becomes

$$\nu = \nu_0\left[1 + \frac{5}{32}\,\frac{\varepsilon m}{\nu_0}\,\frac{A^4}{\pi m \omega_0}\right] = \nu_0 + \frac{5}{64\pi^2}\,\frac{\varepsilon A^4}{\nu_0}$$

or

$$\Delta\nu = \nu - \nu_0 = \frac{5}{64\pi^2}\,\frac{\varepsilon A^4}{\nu_0} \,, \tag{9.34}$$

$$\frac{\Delta\nu}{\nu_0} = \frac{5}{64\pi^2}\,\frac{\varepsilon A^4}{\nu_0^2} \,; \qquad \frac{\Delta\omega}{\omega_0} = \frac{5}{16}\,\frac{\varepsilon A^4}{w_0^2} \,. \tag{9.35}$$

A further example with one degree of freedom is the plane mathematical pendulum with small amplitude. If l is the length of the pendulum and the origin of the coordinate system is assumed to be in the suspension point, then the Hamiltonian reads:

$$\begin{aligned} H &= \frac{p^2}{2ml^2} + mgl(1-\cos\varphi) \\ &\cong \frac{p^2}{2ml^2} + mgl\left(\frac{\varphi^2}{2} - \frac{\varphi^4}{24}\right) . \end{aligned} \tag{9.36}$$

Introducing $I = ml^2$, $\omega_0 = \sqrt{g/l}$ we have

$$H = \frac{p^2}{2I} + \frac{I\omega_0^2}{2}\varphi^2 - \frac{1}{24}I\omega_0^2\varphi^4 = H_0(\mathrm{H.O.}) + \varepsilon H_1 . \tag{9.37}$$

We now substitute $m \to I$ and $q \to \varphi$ in (9.29):

$$\varphi = \sqrt{\frac{J_0}{I\pi\omega_0}}\sin(2\pi w_0) , \quad p = \sqrt{\frac{I\omega_0 J_0}{\pi}}\cos(2\pi w_0) . \tag{9.38}$$

Now we can express H in terms of action-angle variables and in this manner gain access to a perturbative treatment:

$$H = \frac{\omega_0}{2\pi}J_0 - \frac{1}{24}\frac{J_0^2}{I\pi^2}\sin^4(2\pi w_0) , \tag{9.39}$$

with

$$\varepsilon H_1 = -\frac{J_0^2}{24I\pi^2}\sin^4(2\pi w_0) . \tag{9.40}$$

For ε we choose φ_1^2, the maximum angle of the harmonically swinging pendulum (with small amplitude). Then (9.23) tells us that

$$E_1(J) = \overline{H_1(w_0, J)} = -\frac{J^2}{24I\pi^2\varphi_1^2}\overline{\sin^4(2\pi w_0)} .$$

Here, we have

$$\overline{\sin^4(2\pi w_0)} = \int_0^1 dw_0 \sin^4(2\pi w_0) = \frac{3}{8}$$

since

$$\begin{aligned} \sin^4\alpha &= \left(\frac{1}{2\mathrm{i}}\right)^4 \left(\mathrm{e}^{\mathrm{i}\alpha} - \mathrm{e}^{-\mathrm{i}\alpha}\right)^4 = \frac{1}{16}\left[\mathrm{e}^{4\mathrm{i}\alpha} - 6\mathrm{e}^{2\mathrm{i}\alpha} + 6 - 6\mathrm{e}^{-2\mathrm{i}\alpha} + \mathrm{e}^{-4\mathrm{i}\alpha}\right] \\ &= \frac{1}{8}[\cos(4\alpha) - 6\cos(2\alpha) + 3] \end{aligned}$$

so that

$$\overline{\sin^4\alpha} = \tfrac{3}{8} .$$

Up until now we have

$$E_1(J) = -\frac{J^2}{64I\pi^2\varphi_1^2} \; .$$

The frequency change results from this as

$$\Delta\nu = \varepsilon\frac{\partial E_1(J)}{\partial J} = -\frac{J}{32I\pi^2} \; . \tag{9.41}$$

Since we are determining $\Delta\nu$ in first order, we can replace J by J_0 here: $J_0 = (2\pi/\omega_0)E_0$ with $E_0 = I\omega_0^2\varphi_1^2/2$. Then J_0 becomes

$$J_0 = \pi I\omega_0\varphi_1^2 = 2\pi^2\varphi_1^2\nu_0 I \; .$$

We insert this into (9.41) and get

$$\Delta\nu = -\frac{2\pi^2 I\varphi_1^2\nu_0}{32I\pi^2} = -\frac{\varphi_1^2}{16}\nu_0 \tag{9.42}$$

or

$$\frac{\Delta\nu}{\nu_0} = \frac{\nu - \nu_0}{\nu_0} = -\frac{\varphi_1^2}{16} \; . \tag{9.43}$$

13. Superconvergent Perturbation Theory, KAM Theorem (Introduction)

Here we are dealing with an especially fast converging perturbation series, which is of particular importance for the proof of the KAM theorem (cf. below).

Until now we have transformed the Hamiltonian $H = H_0 + \varepsilon H_1$ by successive canonical transformations in such a manner that the order of the perturbation grows by one power in ε with every step. After the nth transformation we therefore obtain

$$\varepsilon H_1 \to \varepsilon^2 H_2 \to \ \dots \ \to \varepsilon^n H_n \ . \tag{13.1}$$

Following Kolmogorov, we can find a succession of canonical transformations for which the order of the perturbation series increases much faster:

$$\varepsilon H_1 \to \varepsilon^2 H_2 \to \varepsilon^4 H_3 \to \ \dots \ \to \varepsilon^{2^{n-1}} H_n \ . \tag{13.2}$$

We should now like to establish an analogy between the two procedures. It is based on the two following methods of finding the zero of a function $f(x)$. We begin by assuming that the zero is at x_0 (unperturbed value of the action J_0). The next improved approximation x_1 is obtained from a Taylor expansion around x_0:

$$f(x) = f\big(x_0\big) + \left.\frac{df(x)}{dx}\right|_{x=x_0} \big(x - x_0\big) + \text{rem.} = 0 \ .$$

If we neglect the remainder, then we obtain as our first approximation

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \ . \tag{13.3}$$

In order to establish the error, we consider the first neglected term in the Taylor series: for this reason, let us define $\varepsilon := (x - x_0)$, and write

$$f(x) = f\big(x_0\big) + f'\big(x_0\big)\varepsilon + \tfrac{1}{2} f''\big(x_0\big)\varepsilon^2 = 0 \ .$$

If we subtract from this expression $f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) = 0$, then we get as error

$$e_1 := x - x_1 = -\frac{1}{2!}\frac{f''(x_0)}{f'(x_0)}\varepsilon^2 \ . \tag{13.4}$$

If we are considering n terms, then we would have to solve the following polynomial (of nth degree) for x_n, in order to determine x_n:

$$\sum_{m=0}^{n} \frac{1}{m!} f^{(m)}(x_0)(x_n - x_0)^m = 0 .$$

If we were to now subtract this result from the Taylor series around x_0, we would obtain, after the nth step, an error of

$$e_n := x - x_n \sim \frac{1}{(n+1)!} \frac{f^{(n+1)}(x_0)}{f'(x_0)} \varepsilon^{n+1} , \tag{13.5}$$

i.e.,

$$e_n \sim \varepsilon^{n+1} . \tag{13.6}$$

One should note that in the denominator, the derivative is always taken at x_0.

Matters look completely different when using Newton's procedure. The first step is identical to that of the foregoing procedure and gives

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{13.7}$$

with the error

$$e_1 := x - x_1 =: \alpha(x_0)\varepsilon^2 ,$$

where

$$\alpha(x_0) = -\frac{1}{2} \frac{f''(x_0)}{f'(x_0)} . \tag{13.8}$$

However, the second step consists of an expansion around x_1 (not x_0!) which we have just found:

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \ldots = 0 .$$

Thus, we obtain as solution for x_2 [unlike in (13.3)]:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} .$$

The error in the second step can now be determined by subtracting the equation

$$f(x_1) + f'(x_1)(x_2 - x_1) = 0$$

from

$$f(x_1) + f'(x_1)(x - x_1) + \frac{1}{2!} f''(x_1)(x - x_1)^2 = 0 .$$

Thus

$$x - x_2 =: e_2 = -\frac{1}{2!}\frac{f''(x_1)}{f'(x_1)}(x - x_1)^2$$
$$= \alpha(x_1)\left(\alpha(x_0)\varepsilon^2\right)^2 = \alpha(x_1)\alpha(x_0)^2\varepsilon^4 \,. \tag{13.9}$$

Hence, it follows for the error e_n after the nth step:

$$e_n \sim \prod_{i=1}^{n}\left(\alpha^{2^{n-i}}(x_{i-1})\right)\varepsilon^{2^n} \,. \tag{13.10}$$

The fast convergence of the Newtonian iteration procedure is evident if we write

$$e_1 \sim \varepsilon^2 \,, \quad e_2 \sim \varepsilon^4 \,, \quad e_3 \sim \varepsilon^8 \,, \quad e_4 \sim \varepsilon^{16} \,, \ldots \,. \tag{13.11}$$

The reason for this fast convergence (superconvergence) is the fact that at each step, $f(x)$ is taken at the just previously calculated approximation x_n, rather than at x_0.

This is precisely the procedure used to prove the KAM theorem. Each new "torus" which was generated by the preceding approximation becomes the basis of the next approximation itself. Thus we do not generate all successive approximations – as in the canonical perturbation series –, always starting with the unperturbed torus (J_i, θ_i), with the Hamiltonian $H_0(J_i)$.

We now look a bit ahead: the crux of the KAM theorem (according to Kolmogorov, Arnol'd and Moser) is that the process of generating "perturbed tori" indeed "almost always" converges for small but finite ε. Thus most of the phase space trajectories remain for all times on tori $\mathcal{M}$ of N dimensions and do not migrate into the entire $2N - 1$ dimensional energy hyperplane. But "almost all" those unperturbed tori found in the proximity of tori whose orbits are closed will be destroyed. These orbits lie on tori with commensurate frequencies:

$$\boldsymbol{\omega}_0 \cdot \boldsymbol{m} = 0 \,. \tag{13.12}$$

These destroyed tori are precisely the ones which give rise to the famous small denominators. But we have already seen that for every arbitrary $\boldsymbol{\omega}_0$, there are "rational tori" which satisfy (13.12). So if we destroy all rational tori and their close neighbors, are there any at all which remain intact – although somewhat deformed? Indeed! In order to understand this, and to specify the width of the destroyed regions, we must concern ourselves briefly with rational and irrational numbers. These are necessary for the arithmetic of torus destruction.

In the following we consider two degrees of freedom. This is the most simple nontrivial case. Then for closed orbits it holds for the frequency ratio that

$$\frac{\omega_{01}}{\omega_{02}} =: \sigma = \frac{r}{s} \,. \tag{13.13}$$

r and s are integers, and σ is therefore rational. A torus with incommensurate frequencies possesses irrational σ and cannot be represented in the form of (13.13).

But one can approximate its frequency ratio arbitrarily precisely by rational σ's. Let us take, for example, the number π:

$$\sigma = \pi = 3.1415926535\ldots$$

with

$$\frac{r}{s} = \frac{3}{1}, \frac{31}{10}, \frac{314}{100}, \frac{3142}{1000}, \frac{31426}{10000}, \ldots .$$

The better approximations contain larger values for r and s. In fact, for each of these approximations, it holds that

$$\left|\sigma - \frac{r}{s}\right| < \frac{1}{s} . \tag{13.14}$$

But we can approximate irrational tori even better (faster): namely, with the help of continued fractions. Here are a few examples:

$$\frac{747}{61} = 12 + \frac{15}{61} = 12 + \frac{1}{61/15} = 12 + \frac{1}{4 + 1/15}$$

$$\frac{7}{10} = \frac{1}{10/7} = \frac{1}{1 + 3/7} = \frac{1}{1 + \frac{1}{7/3}} = \frac{1}{1 + \frac{1}{2+1/3}} .$$

For the number π, matters are more complicated:

$$\begin{aligned}\pi \approx 3.1415926534 &= 3 + \frac{141592654}{10^9} \\ &= 3 + \frac{1}{\frac{10^9}{141592654}} = 3 + \frac{1}{7 + \frac{8851436}{141592654}} \\ &= 3 + \frac{1}{7 + \frac{1}{141592654/8851436}} = 3 + \frac{1}{7 + \frac{1}{15+8821114/8851436}} \\ &= 3 + \frac{1}{7 + \frac{1}{15+\frac{1}{8851436/8821114}}} = 3 + \frac{1}{7 + \frac{1}{15+\frac{1}{1+30322/8821114}}}\end{aligned}$$

$$\pi = 3 + \frac{1}{7 + \frac{1}{15+\frac{1}{1+1/291+\ldots}}} .$$

Thus the approximands of the continued fractions read

$$\pi = \sigma = \frac{r}{s} = \frac{3}{1}, \underbrace{\frac{22}{7}}_{3.14}, \underbrace{\frac{333}{106}}_{3.141}, \underbrace{\frac{355}{113}}_{3.141592}, \ldots .$$

The rational numbers r_n/s_n appearing here are alternatively larger and smaller than σ and approximate σ with quadratic convergence:

$$\left|\sigma - \frac{r_n}{s_n}\right| < \frac{1}{s_n s_{n-1}} \, . \tag{13.15}$$

The slowest convergence, i.e., the most irrational number which one is least able to approximate using continued fractions, is given by the golden mean:

$$\sigma = \frac{1}{1 + \frac{1}{1+1/1+\ldots}} = 0.618033989 \tag{13.16}$$

$$= \frac{\sqrt{5}-1}{2} = \text{“golden mean.''} \tag{13.17}$$

Apparently σ satisfies the equation

$$\sigma = \frac{1}{1+\sigma} \tag{13.18}$$

of which (13.16) is the iteration. Let us write (13.18) in the form

$$\sigma(1+\sigma) = 1 : \;\; \sigma^2 + \sigma - 1 = 0 \, .$$

Then one solution is indeed

$$\sigma = \tfrac{1}{2}(\sqrt{5} - 1) \, .$$

Another famous irrational number is e $= 2.7182818285\ldots$

$$\mathrm{e} = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{1 + \ldots}}}}}} \, .$$

Having completed these mathematical preliminaries, let us return to physics. We know that a system with rational frequency ratios is not integrable – perturbatively speaking. It looks as if the system would at the most be integrable for irrational values of ω_1/ω_2, and that convergence of the perturbation series in ε would exist in this case. We shall therefore first answer the question as to what happens to an integrable unperturbed system $H_0(J_1, J_2)$ whose unperturbed frequency ratio ω_1/ω_2 lies in the neighborhood of an irrational number, and which is perturbed by εH_1. What happens to the rational ω_1/ω_2 in the case of a perturbation will be answered later on.

The KAM theorem now says that if, in addition to other assumptions (cf. below), in particular the functional determinant of the (action-dependent) frequencies does not vanish,

$$\left|\frac{\partial \omega_i}{\partial J_j}\right| \neq 0 \, , \tag{13.19}$$

for those tori, whose frequency ratio ω_1/ω_2 is “sufficiently” irrational,

$$\left|\frac{\omega_1}{\omega_2} - \frac{r}{s}\right| > \frac{K(\varepsilon)}{s^{2.5}} , \tag{13.20}$$

with r and s relatively prime, the iterated (according to Kolmogorov) perturbation series for the generator $W(\theta_k^0, J_k)$ converges (for small enough ε), and therefore the invariant tori are not destroyed.

One should note that the set of frequency ratios or space for the KAM curves, i.e., curves for which condition (13.20) holds, indeed amounts to a finite part of, e.g., the interval $0 \leqslant \omega_1/\omega_2 \leqslant 1$, because we obtain for the total length L of the intervals for which (13.20) is not valid, i.e., $|\omega_1/\omega_2 - r/s| < K(\varepsilon)/s^{2.5}$,

$$L < \sum_{s=1}^{\infty} s\frac{K(\varepsilon)}{s^{2.5}} = K(\varepsilon)\zeta(1.5) \approx 2.6K(\varepsilon) < 1 . \tag{13.21}$$

Here, $K(\varepsilon)/s^{2.5}$ is the width of an interval around the rational value r/s, for which (13.20) does not hold, and s is the number of r-values with $r/s < 1$.

0	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{4}{5}$	1	
	$\frac{K}{55.9}$	$\frac{K}{32}$	$\frac{K}{15.6}$	$\frac{K}{55.9}$	$\frac{K}{5.7}$	$\frac{K}{55.9}$	$\frac{K}{15.6}$	$\frac{K}{32}$	$\frac{K}{55.9}$		$\frac{\omega_1}{\omega_2}$

For sufficiently large ε, εH_1 destroys all tori. The last KAM torus which is destroyed is the one whose frequency ratio is the most irrational; namely,

$$\frac{\omega_1}{\omega_2} = \frac{1}{2}(\sqrt{5} - 1) .$$

By way of illustration of the KAM theorem, we consider a system from the dynamics of the solar system. (Admittedly, the following treatment will be somewhat oversimplified; nevertheless, it contains much truth.) Let three bodies move under the influence of mutual gravitation, e.g., a very massive main body M (sun or Saturn), a perturbing body m (Jupiter or moon of Saturn) and a test body of mass μ, whose long-term behavior we wish to study. Furthermore, let us assume that $\mu \ll m \ll M$; second, that all these bodies are moving in a fixed plane; and third, that m is rotating on a circle (rather than on an ellipse) around the common center of mass of M and m.

The Hamiltonian for the motion of the test body μ in presence of the gravitational fields of m and M is thus

$$H(\boldsymbol{q}, \boldsymbol{p}; t) = \frac{\boldsymbol{p}^2}{2\mu} - \frac{GM\mu}{r} - \frac{Gm\mu}{|\boldsymbol{q} - \boldsymbol{r}_m(t)|} . \tag{13.22}$$

It is obvious that (13.22) is a time-dependent Hamiltonian. Now, according to our assumption, m is supposed to move around the center of mass at the angular velocity Ω, so that it seems logical to eliminate the time dependence of H by making a transformation to the comoving, rotating system. In this system, m is at rest. Then the new conservative Hamiltonian reads

$$H(\boldsymbol{q}, \boldsymbol{p}) = \frac{\boldsymbol{p}^2}{2\mu} - \Omega p_\phi - \frac{GM\mu}{r} - \frac{Gm\mu}{|\boldsymbol{q} - \boldsymbol{r}_m|} \tag{13.23}$$

$$= H_0(\boldsymbol{q}, \boldsymbol{p}) + \varepsilon H_1 , \tag{13.24}$$

where

$$H_{m=0} \equiv H_0(\boldsymbol{q} \equiv \boldsymbol{r}, \boldsymbol{p}) = \frac{1}{2\mu}\left(p_r^2 + \frac{1}{r^2}p_\phi^2\right) - \Omega p_\phi - \frac{GM\mu}{r} \tag{13.25}$$

and

$$\varepsilon H_1 = -\frac{Gm\mu}{|\boldsymbol{q} - \boldsymbol{r}_m|} . \tag{13.26}$$

Note that we have replaced $\boldsymbol{q} \to \boldsymbol{r}$ as $m \to 0$ in (13.25). Neither t nor ϕ appears in (13.25), so that the two constants of motion are p_ϕ and H_0. Equation (13.26) gives the nonintegrable perturbation term in which m plays the role of the small parameter.

Specifying the conserved quantities p_ϕ and H_0, then a certain torus is defined. The action variables J_ϕ and J_r in terms of p_ϕ and H_0 are given as follows:

$$J_\phi = \frac{1}{2\pi}\int_0^{2\pi} p_\phi \, d\phi = p_\phi \tag{13.27}$$

$$J_r = \frac{1}{2\pi}\oint p_r \, dr = \frac{1}{2\pi}\oint dr \sqrt{2\mu\left(H_0 - \Omega J_\phi + \frac{GM\mu}{r}\right) - \frac{J_\phi^2}{r^2}} \tag{13.28}$$

$$= -J_\phi + \frac{GM\mu^2}{\sqrt{-2\mu(H_0 + \Omega J_\phi)}} . \tag{13.29}$$

The new Hamiltonian – relative to the rotating system – thus reads, as a function of the action variables J_ϕ and J_r:

$$H_0\big(J_r, J_\phi\big) = -\Omega J_\phi - \frac{G^2M^2\mu^3}{2(J_r + J_\phi)^2} . \tag{13.30}$$

Hence, the unperturbed Hamiltonian is a nonlinear function of the actions. For the unperturbed frequencies we find, using $\omega_{0i} = \partial H_0/\partial J_i$:

$$\omega_{0r} = \frac{G^2M^2\mu^3}{(J_r + J_\phi)^3} , \quad \omega_{0\phi} = -\Omega + \frac{G^2M^2\mu^3}{(J_r + J_\phi)^3} . \tag{13.31}$$

Here,

$$\omega_\mu := \frac{G^2M^2\mu^3}{(J_r + J_\phi)^3} \tag{13.32}$$

is the frequency of the Kepler motion relative to the nonrotating coordinate system in which the r- and ϕ-motion have the same frequency (accidental degeneracy in the $1/r$-potential). Then we finally obtain

$$\omega_{0r} = \omega_\mu \, , \quad \omega_{0\phi} = -\Omega + \omega_\mu \, . \tag{13.33}$$

So the decision as to the regular or stochastic behavior of the motion of the problem perturbed by εH_1 depends on the following frequency ratio:

$$\frac{\omega_{0\phi}}{\omega_{0r}} = 1 - \frac{\Omega}{\omega_\mu} \, . \tag{13.34}$$

The invariant tori are thus destroyed if the frequency ratio, Ω/ω_μ, of the m- and μ-motion is rational. In fact, there are distributions of test bodies in the solar system in which gaps between tori can be observed. This is the case for the asteroid belt between Mars and Jupiter. Here, the sun is the main body, and Jupiter, the perturbing body. The test mass μ is any asteroid. According to the KAM theorem, one should expect gaps (instabilities) in the asteroid belt if the frequency of the asteroids and the Jupiter frequency Ω_J are commensurate. These gaps were observed by Kirkwood in 1866 and are therefore called Kirkwood gaps. They occur at $\omega_\mu/\Omega_J = 2, 3, 4$ especially clearly, at $\omega_\mu/\Omega_J = 3/2, 5/2, 7/2$, less so.