Chapter 2
Theoretical Fundamentals

In this chapter, we summarize some theoretical fundamentals. We assume that the reader is already familiar with these basic facts. The main purpose of this chapter is to introduce the notation that is used in this book and to provide a reference. Therefore, the explanations are brief, and no proofs are given.

2.1 Fourier Analysis and Application to Beam Signals

In this section, several formulas for Fourier series and the Fourier transform are summarized. However, we do not discuss the properties of a function that are necessary for the existence of the transformation. For those foundations, the reader should consult the references cited here.

2.1.1 Fourier Series

A real-valued periodic function $f(t)$ with period T may be decomposed into Fourier components according to the **Fourier series**

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \, e^{jn\omega t} \qquad \text{with } \omega = \frac{2\pi}{T},$$ (2.1)

where the complex coefficients c_n are determined by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varphi) \, e^{-jn\varphi} \, \mathrm{d}\varphi$$

or by

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t)\, e^{-jn\omega t}\, dt, \tag{2.2}$$

where we made the substitution $\varphi = \omega t$.

With the substitution $x = t + T$, we obtain

$$\int_{-T/2}^{0} f(t)\, e^{-jn\omega t}\, dt = \int_{T/2}^{T} f(x - T)\, e^{-jn\omega x}\, e^{jn\omega T}\, dx.$$

Due to $\omega = \frac{2\pi}{T}$, the last exponential function equals 1. Furthermore, we have $f(x - T) = f(x)$, so that

$$\int_{-T/2}^{0} f(t)\, e^{-jn\omega t}\, dt = \int_{T/2}^{T} f(x)\, e^{-jn\omega x}\, dx = \int_{T/2}^{T} f(t)\, e^{-jn\omega t}\, dt$$

holds. Therefore, we may use

$$c_n = \frac{1}{T} \int_{0}^{T} f(t)\, e^{-jn\omega t}\, dt \tag{2.3}$$

instead of Eq. (2.2).

2.1.2   Spectrum of a Dirac Comb

In this book, the Dirac delta distribution is used in a heuristic way without the foundations of distribution theory. Therefore, the reader should be aware that the results presented still have to be proven mathematically. For example, we use the formula

$$\int_{-\infty}^{+\infty} f(x)\, \delta(x - x_0)\, dx = f(x_0),$$

even though it does not have any meaning in the scope of classical analysis.

A strongly bunched beam may be approximated by a sum of Dirac delta pulses

$$f(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT),$$

which is called **Dirac comb**. For this special sum of Dirac pulses, one obtains the following Fourier coefficients (only the Dirac pulse with $k = 0$ is located inside the interval $-T/2 \leq t \leq +T/2$):

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} \delta(t)\, e^{-jn\omega t}\, dt = \frac{1}{T}. \tag{2.4}$$

Hence, all coefficients are equal. According to Eq. (2.1), we get

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{jn\omega t}.$$

This can also be written as

$$\sum_{k=-\infty}^{\infty} \omega\, \delta(\omega t - \omega kT) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{jn\omega t}$$

$$\Rightarrow 2\pi \sum_{k=-\infty}^{\infty} \delta(\varphi - 2\pi k) = \sum_{n=-\infty}^{\infty} e^{jn\varphi}.$$

2.1.3   Different Representations of the Fourier Series

The general definition of the Fourier series shows that the c_n are defined in such a way that both positive and negative frequencies occur. If only positive frequencies are to be allowed, one may write Eq. (2.1) as follows:

$$f(t) = c_0 + \sum_{n=1}^{\infty} \left(c_n e^{jn\omega t} + c_{-n} e^{-jn\omega t} \right) = \tag{2.5}$$

$$= c_0 + \sum_{n=1}^{\infty} \left(c_n \left[\cos(n\omega t) + j\, \sin(n\omega t) \right] \right.$$

$$\left. + c_{-n} \left[\cos(n\omega t) - j\, \sin(n\omega t) \right] \right). \tag{2.6}$$

We obtain the result

$$f(t) = c_0 + \sum_{n=1}^{\infty} \left[(c_n + c_{-n}) \cos(n\omega t) + j(c_n - c_{-n}) \sin(n\omega t) \right].$$

By means of the definition

$$a_n = c_n + c_{-n}$$

and

$$b_n = j(c_n - c_{-n}),$$

one obtains

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos(n\omega t) + b_n \sin(n\omega t) \right]. \tag{2.7}$$

Taking $a_0 = 2c_0$ and $b_0 = 0$ into account, one may calculate the coefficients c_n if a_n and b_n are known:

$$c_n = \frac{a_n - jb_n}{2}. \tag{2.8}$$

For the special case that $c_n = 1/T$ holds for all n (Dirac comb; see Sect. 2.1.2), one obtains $a_n = 2/T$ and $b_n = 0$. According to Eq. (2.7), this means that the average, i.e., the DC component, of a strongly bunched beam is exactly one-half the fundamental harmonic:

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \frac{1}{T} + \frac{2}{T} \sum_{n=1}^{\infty} \cos(n\omega t). \tag{2.9}$$

Now we return to the general case. Instead of using a_n and b_n, one may also use amplitudes and phases:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} d_n \, \cos(n\omega t + \varphi_n). \tag{2.10}$$

A comparison with Eq. (2.7) shows that

$$a_n \cos(n\omega t) + b_n \sin(n\omega t) = d_n \, \cos(n\omega t + \varphi_n)$$

$$\Rightarrow a_n \cos(n\omega t) + b_n \sin(n\omega t) = d_n \, \cos(n\omega t) \, \cos \varphi_n - d_n \, \sin(n\omega t) \, \sin \varphi_n.$$

This leads to the following conditions:

$$a_n = d_n \, \cos \varphi_n, \tag{2.11}$$

$$b_n = -d_n \, \sin \varphi_n. \tag{2.12}$$

According to Eq. (2.8), we therefore have

$$\varphi_n = \angle\, c_n \qquad (2.13)$$

and

$$d_n = \sqrt{a_n^2 + b_n^2}. \qquad (2.14)$$

Due to

$$|c_n| = \frac{1}{2}\sqrt{a_n^2 + b_n^2}, \qquad (2.15)$$

one obtains

$$d_n = 2|c_n| \qquad (2.16)$$

as the physical amplitudes (peak values). By inserting Eqs. (2.11) and (2.12) into Eq. (2.8), one gets

$$c_n = \frac{d_n}{2} e^{j\varphi_n}.$$

The same result is obtained by combining Eqs. (2.13)–(2.15).

2.1.4   Discrete Fourier Transform

The discrete Fourier transform is a powerful tool for spectral analysis of signals that are given in digital form, e.g., on a computer. Therefore, we briefly discuss some important features here.

2.1.4.1   Motivation of the Transformation Formula

Let us now assume that a real-valued periodic function $f(t)$ with period $T = \frac{2\pi}{\omega}$ is discretized according to

$$f_k = f(k\Delta t),$$

where k is an integer. The period T is divided into $N \in \mathbb{N}$ time intervals

$$\Delta t = \frac{T}{N}$$

such that $f_0 = f_N$ holds. Therefore, the N samples $f_0, f_1, \ldots, f_{N-1}$ are sufficient to describe the function $f(t)$, provided that N is large enough. We now replace the integral in Eq. (2.3) by the Riemann sum

$$c_n \approx \frac{1}{T} \sum_{k=0}^{N-1} f(k\Delta t)e^{-jn\omega k\Delta t} \Delta t = \frac{1}{T} \sum_{k=0}^{N-1} f_k e^{-j2\pi nk/N} \Delta t = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j2\pi nk/N}.$$

This formula is used to define the **discrete Fourier transform (DFT)**

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k \, e^{-j2\pi nk/N}. \tag{2.17}$$

This obviously yields an approximation of the Fourier coefficients c_n of the periodic function $f(t)$, provided that the number N of samples $x_k = f(k\Delta t)$ is large enough.

2.1.4.2 Symmetry Relations

Based on Eq. (2.17), we find that

$$X_{n+N} = \frac{1}{N} \sum_{k=0}^{N-1} x_k \, e^{-j2\pi nk/N} \, e^{-j2\pi Nk/N} = \frac{1}{N} \sum_{k=0}^{N-1} x_k \, e^{-j2\pi nk/N} \cdot 1 = X_n.$$

Therefore, all X_n are known if those for $0 \leq n \leq N - 1$ are specified. One sees that for a sample $(x_0, x_1, \ldots, x_{N-1})$, one obtains a sample $(X_0, X_1, \ldots, X_{N-1})$ as the spectrum.

Since we have assumed that the signal $f(t)$ is real-valued and periodic, the same is true for the samples x_k. Based on Eq. (2.17), it is then obvious that the symmetry relation

$$X_{-n} = X_n^*$$

holds. We may also combine these two symmetry relations to obtain

$$X_{N-n} = X_{-(n-N)} = X_{n-N}^* = X_n^*.$$

Therefore, only about one-half of the coefficients X_n with $0 \leq n \leq N - 1$ have to be calculated.

Table 2.1 Overview of DFT components of real-valued signals

Time	Sample	Frequency	Spectral component	Comment
0	x_0	0	X_0 (real)	DC component of the signal
Δt	x_1	$\frac{1}{T} = \frac{1}{N\Delta t} = \frac{1}{N} f_{\text{sampl}}$	X_1	Peak value: $2|X_1|$
$2\Delta t$	x_2	$\frac{2}{T} = \frac{2}{N\Delta t} = \frac{2}{N} f_{\text{sampl}}$	X_2	Peak value: $2|X_2|$
$3\Delta t$	x_3	$\frac{3}{T} = \frac{3}{N\Delta t} = \frac{3}{N} f_{\text{sampl}}$	X_3	Peak value: $2|X_3|$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(N-2)\Delta t$	x_{N-2}	$\frac{N-2}{T} = \frac{N-2}{N\Delta t} = \frac{N-2}{N} f_{\text{sampl}}$	$X_{N-2} = X_2^*$	Peak value: $2|X_2|$
$(N-1)\Delta t$	x_{N-1}	$\frac{N-1}{T} = \frac{N-1}{N\Delta t} = \frac{N-1}{N} f_{\text{sampl}}$	$X_{N-1} = X_1^*$	Peak value: $2|X_1|$
$N\Delta t$	x_N	$\frac{N}{T} = \frac{1}{\Delta t} = f_{\text{sampl}}$	$X_N = X_0$	Repetition
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

2.1.4.3 Interpretation of the Spectral Components

According to Eq. (2.1), the sample X_0 belongs to the DC component of the signal. The sample X_1 obviously belongs to the angular frequency

$$1 \cdot \omega = \frac{2\pi}{T}.$$

Therefore, the spectrum $(X_0, X_1, \ldots, X_{N-1})$ has a resolution of $f = 1/T$, where T is the total time that passes between the samples x_0 and x_N. It is obvious that X_{N-1} belongs to the frequency

$$f_{\text{max}} = \frac{N-1}{T} = \frac{N-1}{N} \frac{1}{\Delta t} \approx \frac{1}{\Delta t} = f_{\text{sampl}}.$$

This approximation is, of course, valid only for large samples with $N \gg 1$. Hence we conclude that the frequency resolution is given by the inverse of the total time T, whereas the maximum frequency is determined by the sampling frequency $f_{\text{sampl}} = 1/\Delta t$. However, due to $X_{N-n} = X_n^*$, only one-half of this frequency range between 0 and f_{max} actually contains information. In other words, and in compliance with the **Nyquist–Shannon sampling theorem**, sampling has to take place with at least twice the signal bandwidth.

These properties are visualized in Table 2.1.

If one makes sure that the N equidistant samples x_n of the periodic function represent an integer number of periods (so that duplicating $(x_0, x_1, \ldots, x_{N-1})$ does not introduce any severe discontinuities), one may obtain good results even without sophisticated windowing techniques.

For the interpretation of the spectrum, please note that the DC component is equal to

$$f_{\text{DC}} = \frac{a_0}{2} = c_0 = X_0,$$

i.e., to the first value of the DFT.

According to Eq. (2.16), the amplitude (peak value) at the frequency p/T is given by

$$d_p = 2|c_p| = 2|X_p|.$$

The discussion above shows that the sample $(X_0, X_1, \ldots, X_{N-1})$ contains all the information about the spectrum, but that the DFT spectrum is infinite. It does not even decrease with increasing frequencies. At first glance, this looks strange, but in our introduction to the DFT, we assumed only that the integral over Δt may approximately be replaced by a product with Δt. We made no assumption as to how the function $f(t)$ varies in the interval Δt. This explains the occurrence of the high-frequency components.

It should be clear from the Nyquist–Shannon sampling theorem that the spectrum for frequencies larger than $f_{max}/2$ cannot contain any relevant information, since the sampling frequency is fixed at $\Delta t \approx 1/f_{max}$.

Therefore, in the next section, we filter out those frequencies to obtain the inverse transform.

2.1.4.4 Inverse DFT

As mentioned above, the Nyquist–Shannon sampling theorem tells us that we should consider only frequencies f_x with

$$-\frac{f_{max}}{2} \leq f_x \leq +\frac{f_{max}}{2}.$$

This corresponds to

$$-\frac{N-1}{2T} \leq f_x \leq +\frac{N-1}{2T},$$

or

$$-\frac{N-1}{2}\omega \leq \omega_x \leq +\frac{N-1}{2}\omega.$$

For the sake of simplicity, we assume that $N \geq 3$ is an odd number. If we have a look at Eq. (2.1),

$$f(t) = \sum_{n=-\infty}^{\infty} c_n\, e^{jn\omega t},$$

it becomes clear that only those n with

$$-\frac{N-1}{2} \leq n \leq +\frac{N-1}{2}$$

lead to the aforementioned frequencies $\omega_x = 2\pi f_x = n\omega$. Therefore, we expect to be able to reconstruct the signal based on

$$f(t) = \sum_{n=-(N-1)/2}^{+(N-1)/2} c_n \, e^{jn\omega t}.$$

We now apply the discretization

$$f_k = f(k\Delta t) = \sum_{n=-(N-1)/2}^{+(N-1)/2} c_n \, e^{jn\omega k \Delta t} = \sum_{n=-(N-1)/2}^{+(N-1)/2} c_n \, e^{j2\pi nk/N} \qquad (2.18)$$

and obtain

$$\sum_{n=(N+1)/2}^{N-1} c_n \, e^{j2\pi nk/N} = \sum_{l=-N/2+1/2}^{-1} c_{l+N} \, e^{j2\pi k \frac{l+N}{N}}.$$

Here we introduced the new summation index $l = n - N$. The last formula leads to

$$\sum_{n=(N+1)/2}^{N-1} c_n \, e^{j2\pi nk/N} = \sum_{l=-(N-1)/2}^{-1} c_l \, e^{j2\pi kl/N}.$$

On the right-hand side, we may now rename l as n again. This shows that the sum from $-(N-1)/2$ to -1 included in Eq. (2.18) may be replaced by the sum from $(N+1)/2$ to $N-1$:

$$f_k = \sum_{n=0}^{N-1} c_n \, e^{j2\pi nk/N}.$$

This defines the formula for the inverse DFT (not only for odd N):

$$\boxed{x_k = \sum_{n=0}^{N-1} X_n \, e^{j2\pi nk/N}.}$$

Please note that in the literature, the factor $1/N$ is sometimes not included in the definition of the DFT, but it appears in that of the inverse DFT. Our choice was determined by the close relationship to the Fourier series coefficients discussed above. Apart from the factor $1/N$, the DFT and the inverse DFT differ only by the sign in the argument of the exponential function.

2.1.4.5   Conclusion

We have summarized only a few basic facts that will help the reader to interpret the DFT correctly. There are many other properties that cannot be mentioned here.

For large sample sizes equal to a power of 2, the so-called **fast Fourier transform (FFT)** algorithm may be used, which is a dramatically less time-consuming implementation of the DFT.

2.1.5   Fourier Transform

The **Fourier transform** $X(\omega)$ of a real-valued function $x(t)$ depending on the time variable t is given by

$$X(\omega) = \int_{-\infty}^{+\infty} x(t)\, e^{-j\omega t}\, \mathrm{d}t, \tag{2.19}$$

the inverse transform by

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega)\, e^{j\omega t}\, \mathrm{d}\omega. \tag{2.20}$$

This relation is visualized by the correspondence symbol

$$x(t) \circ\!\!\!-\!\!\!-\!\!\bullet X(\omega).$$

The Fourier transform is a linear transformation. It is used to determine the frequency spectrum of signals, i.e., it transforms the signal $x(t)$ from the time domain into the frequency domain. It is possible to generalize the definition of the Fourier transform to generalized functions (i.e., distributions), which also include the Dirac function [1, 2].

Please note that various definitions for the Fourier transform and for its inverse transform exist in the literature. The factor $\frac{1}{2\pi}$ may be distributed among the original transformation and the inverse transformation in a different way, and even the sign of the argument of the exponential function may be defined in the opposite way.

Some common Fourier transforms are summarized in Table A.3 on p. 417. Further relations can also be found using symmetry properties of the Fourier transform. Consider the Fourier transform

$$x(t) \quad \circ\!\!\!-\!\!\!-\!\!\bullet \quad X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, \mathrm{d}t.$$

If the time t in $x(t)$ is replaced by ω, and $x(\omega)$ is regarded as a Fourier transform, its inverse transform is given by

$$x(\omega) \quad \bullet\!\!-\!\!-\!\!\circ \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} x(\omega) \, e^{j\omega t} \, d\omega = \frac{1}{2\pi} X(-t).$$

In other words, the inverse transform of $x(\omega)$ is obtained by replacing ω in the function $X(\omega)$ by $-t$.

2.1.5.1 Fourier Transform of a Single Cosine Pulse

Let

$$x(t) = \begin{cases} 1 + \cos(\Omega t) & \text{for } -\pi < \Omega t < \pi, \\ 0 & \text{otherwise,} \end{cases} \tag{2.21}$$

define a single cosine pulse. This leads to

$$X(\omega) = \int_{-\infty}^{+\infty} x(t) \, e^{-j\omega t} \, dt = \int_{-\pi/\Omega}^{+\pi/\Omega} [1 + \cos(\Omega t)] \, e^{-j\omega t} \, dt =$$

$$= \int_{-\pi/\Omega}^{+\pi/\Omega} \left[e^{-j\omega t} + \frac{1}{2} e^{j(\Omega-\omega)t} + \frac{1}{2} e^{-j(\Omega+\omega)t} \right] dt =$$

$$= \left[\frac{e^{-j\omega t}}{-j\omega} + \frac{1}{2} \frac{e^{j(\Omega-\omega)t}}{j(\Omega-\omega)} + \frac{1}{2} \frac{e^{-j(\Omega+\omega)t}}{-j(\Omega+\omega)} \right]_{-\pi/\Omega}^{+\pi/\Omega} =$$

$$= \sin\left(\pi \frac{\omega}{\Omega}\right) \left[\frac{2}{\omega} + \frac{1}{\Omega-\omega} - \frac{1}{\Omega+\omega} \right] = \sin\left(\pi \frac{\omega}{\Omega}\right) \frac{2\Omega^2}{\omega(\Omega^2 - \omega^2)}$$

$$\Rightarrow X(\omega) = \frac{2\pi}{\Omega} \frac{\text{si}\left(\pi \frac{\omega}{\Omega}\right)}{1 - \left(\frac{\omega}{\Omega}\right)^2}. \tag{2.22}$$

In the last equation, we used the definition

$$\text{si}(x) = \begin{cases} \frac{\sin x}{x} & \text{for } x \neq 0, \\ 1 & \text{for } x = 0. \end{cases}$$

For the sake of uniqueness, we call this function $\text{si}(x)$ instead of $\text{sinc}(x)$.

2.1.5.2 Convolution

The **convolution** is given by

$$\boxed{h(t) * x(t) = \int_{-\infty}^{+\infty} h(\tau) \, x(t - \tau) \, d\tau,}$$

and one obtains

$$x(t) * h(t) = h(t) * x(t) \circ\!\!-\!\!\bullet H(\omega)X(\omega).$$

We consider the special case that

$$h(t) = \sum_k \delta(t - T_k)$$

is a sequence of Dirac pulses. This leads to

$$h(t) * x(t) = \sum_k \int_{-\infty}^{+\infty} \delta(\tau - T_k)\, x(t - \tau)\, d\tau = \sum_k x(t - T_k).$$

Hence, by convolution with a sequence of Dirac pulses, we may produce a repetition of the function $x(t)$ at the locations of the delta pulses.

2.1.5.3   Relation to the Fourier Series

We consider the special case

$$X(\omega) = \sum_{k=-\infty}^{+\infty} p_k\, \delta(\omega - k\omega_0).$$

According to Eq. (2.20), this leads to

$$x(t) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_k\, \delta(\omega - k\omega_0)\, e^{j\omega t}\, d\omega = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} p_k\, e^{jk\omega_0 t}.$$

If we set

$$p_k = 2\pi c_k,$$

we obtain the correspondence

$$x(t) = \sum_{k=-\infty}^{+\infty} c_k\, e^{jk\omega_0 t} \qquad \circ\!\!-\!\!\bullet \qquad X(\omega) = 2\pi \sum_{k=-\infty}^{+\infty} c_k\, \delta(\omega - k\omega_0),$$

which is an ordinary Fourier series, as Eq. (2.1) shows.

   Hence, if we calculate the Fourier transform of a periodic function with period $T_0 = \frac{2\pi}{\omega_0}$, we get a sum of Dirac pulses that are multiplied by 2π and the Fourier coefficients. The factor 2π is obvious because of the correspondence

$$1 \circ\!\!-\!\!\bullet 2\pi\delta(\omega).$$

2.1.6 Consequences for the Spectrum of the Beam Signal

We first model an idealized beam signal $h(t)$ as a periodic sequence of Dirac pulses. Even if the bunches oscillate in the longitudinal direction, periodicity may be satisfied if the beam signal repeats itself after one synchrotron oscillation period. The sequence of delta pulses will be defined by

$$h(t) = \sum_k \delta(t - T_k)$$

as above. Thus, we get a realistic beam signal by convolution with the time function $x(t)$, which represents a single bunch:

$$y(t) = h(t) * x(t).$$

Since $h(t)$ is to be periodic, it may be represented by a Fourier series. As shown in the previous section, this leads to the Fourier transform

$$H(\omega) = 2\pi \sum_{k=-\infty}^{+\infty} c_k^{h(t)} \, \delta(\omega - k\omega_0).$$

The function $x(t)$ describes a single pulse and is therefore equal to zero outside a finite interval. Therefore, the spectrum $X(\omega)$ will be continuous. This shows that

$$Y(\omega) = H(\omega) \, X(\omega)$$

$$= 2\pi \sum_{k=-\infty}^{+\infty} c_k^{h(t)} \, X(\omega) \, \delta(\omega - k\omega_0)$$

$$= 2\pi \sum_{k=-\infty}^{+\infty} c_k^{h(t)} \, X(k\omega_0) \, \delta(\omega - k\omega_0)$$

is a Fourier series whose Fourier coefficients are

$$c_k^{y(t)} = c_k^{h(t)} \, X(k\omega_0). \tag{2.23}$$

_____

As an example and as a test of the results obtained so far, we analyze the convolution of a Dirac comb

$$h(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_0)$$

with a single cosine pulse. According to Eq. (2.4), the Fourier coefficients of the Dirac comb are

$$c_k^{h(t)} = \frac{1}{T_0}.$$

Here T_0 denotes the time span between the pulses. For the single cosine pulse with time span $T = \frac{2\pi}{\Omega}$ that was defined in Eq. (2.21), one obtains—based on Eq. (2.22)—the Fourier transform

$$X(\omega) = \frac{2\pi}{\Omega} \frac{\text{si}\left(\pi \frac{\omega}{\Omega}\right)}{1 - \left(\frac{\omega}{\Omega}\right)^2}.$$

According to Eq. (2.23), the Fourier coefficients of the convolution function $y(t) = h(t) * x(t)$ are therefore

$$c_k^{y(t)} = \frac{1}{T_0} \frac{2\pi}{\Omega} \frac{\text{si}\left(\pi k \frac{\omega_0}{\Omega}\right)}{1 - \left(k \frac{\omega_0}{\Omega}\right)^2} = \frac{\omega_0}{\Omega} \frac{\text{si}\left(\pi k \frac{\omega_0}{\Omega}\right)}{1 - \left(k \frac{\omega_0}{\Omega}\right)^2}. \tag{2.24}$$

We will now analyze this result for several special cases.

- **Constant beam current:** In this first case, we assume that the different single-cosine pulses overlap according to $T = 2T_0$, which is equivalent to $\omega_0 = 2\Omega$. In this case, we obtain $c_k^{y(t)} = 0$ for $k \neq 0$. For $c_0^{y(t)}$, which corresponds to the DC component, one obtains

$$c_0^{y(t)} = \frac{1}{T_0} \frac{2\pi}{\Omega} = 2,$$

  which is the expected result for a constant function that equals 2.
- **Continuous sine wave:** In this case, we make use of the simplification $\Omega = \omega_0$, so that $y(t)$ corresponds to a simple cosine function that is shifted upward:

$$c_k^{y(t)} = \frac{\text{si}(\pi k)}{1 - k^2}.$$

We obviously have

$$c_0^{y(t)} = 1.$$

For $k = \pm 1$, we may use l'Hôpital's rule:

$$c_{\pm 1}^{y(t)} = \lim_{k \to \pm 1} \frac{\text{si}(\pi k)}{1 - k^2} = \lim_{k \to \pm 1} \frac{\sin(\pi k)}{\pi \left(k - k^3\right)} = \lim_{k \to \pm 1} \frac{\pi \, \cos(\pi k)}{\pi \left(1 - 3k^2\right)} = \frac{1}{2}.$$

All other coefficients are zero. Thus we obtain

$$y(t) = \sum_{k=-\infty}^{+\infty} c_k^{y(t)} e^{jk\omega_0 t} = 1 + \frac{1}{2} e^{j\omega_0 t} + \frac{1}{2} e^{-j\omega_0 t} = 1 + \cos(\omega_0 t),$$

which is in accordance with our expectation.

- **Dirac comb:** For this last case, we first observe that the area under each single-cosine pulse defined in Eq. (2.21) is T. If we want to have an area of 1 instead, we have to divide the function $y(t)$ by T:

$$\tilde{y}(t) = \frac{y(t)}{T}.$$

Hence, the Fourier coefficients in Eq. (2.24) also have to be divided by T:

$$c_k^{\tilde{y}(t)} = \frac{1}{T_0} \frac{\text{si}\left(\pi k \frac{\omega_0}{\Omega}\right)}{1 - \left(k \frac{\omega_0}{\Omega}\right)^2}.$$

We now consider the case $T \to 0$ while assuming a fixed value of T_0. Hence $\omega_0/\Omega \to 0$, and we obtain

$$c_k^{\tilde{y}(t)} = \frac{1}{T_0},$$

which is the expected result for a Dirac comb.

Finally, our simple beam signal model that was constructed by a combination of single-cosine pulses is able to describe all states between unbunched beams and strongly bunched beams. In the case of long bunches (continuous sine wave), the DC current equals the RF current amplitude. As the bunches become shorter ($\omega_0 < \Omega$), Eq. (2.24) can be used to determine the ratio between RF current amplitude and DC current.

2.2 Laplace Transform

The **Laplace transform** is one of the standard tools used to analyze closed-loop control systems. In the scope of the book at hand, we deal only with the one-sided Laplace transform [3, 4], which is useful because processes can be described whereby signals are switched on at $t = 0$. Hence, the name "Laplace transform" will be used as a synonym for "one-sided Laplace transform." Such a one-sided Laplace transform of a function $f(t)$ with $f(t) = 0$ for $t < 0$ is given by

$$\boxed{F(s) = \int_0^\infty f(t)\, e^{-st}\, \mathrm{d}t.} \tag{2.25}$$

Here $s = \sigma + j\omega$ is a complex parameter. It is obvious that the Laplace transform has a close relationship to the Fourier transform that is obtained for $\sigma = 0$ if only functions with $f(t) = 0$ for $t < 0$ are allowed. The real part of s is usually introduced to obtain convergence for a larger class of functions (please note that the Fourier transform of a sine or cosine function already leads to nonclassical Dirac pulses, as we saw in Sect. 2.1.5.3).

The Laplace transform $F(s)$ of a function $f(t)$ is an analytic function, and there is a unique correspondence between $f(t)$ and $F(s)$ if the classes of functions/distributions that are considered in the time domain and the Laplace domain are chosen accordingly [1, 4]. Since the integral in Eq. (2.25) exists only in some region of the complex plane, the Laplace transform is initially defined in only this region as well. If, however, a closed-form expression is obtained for the Laplace transform, e.g., a rational function, it is possible to extend the domain of definition by means of analytic continuation (cf. [5, Sect. 2.1]; [6, Sect. 10-9]; [7, Sect. 5.5.4]). Therefore, the Laplace transform $F(s)$ should be defined as the analytic continuation of the function defined by Eq. (2.25). Apart from poles, a Laplace transform $F(s)$ may thus be defined in the whole complex plane.

Like the Fourier transform, the Laplace transform is a linear transformation. If according to

$$f(t) \circ\!\!-\!\!\bullet F(s), \qquad g(t) \circ\!\!-\!\!\bullet G(s),$$

we use the correspondence symbol again, the Laplace transform has the following properties (n is a positive integer, and a is a real number):

- Laplace transform of a derivative[1]:

$$\frac{\mathrm{d}f}{\mathrm{d}t} \circ\!\!-\!\!\bullet s\, F(s) - f(0+),$$

$$\frac{\mathrm{d}^n f}{\mathrm{d}t^n} \circ\!\!-\!\!\bullet s^n F(s) - s^{n-1} f(0+) - s^{n-2}\frac{\mathrm{d}f}{\mathrm{d}t}(0+) - \cdots - s^0 \frac{\mathrm{d}^{n-1} f}{\mathrm{d}t^{n-1}}(0+).$$

- Derivative of a Laplace transform:

$$-t\, f(t) \circ\!\!-\!\!\bullet \frac{\mathrm{d}F(s)}{\mathrm{d}s}, \qquad t^n\, f(t) \circ\!\!-\!\!\bullet (-1)^n \frac{\mathrm{d}^n F(s)}{\mathrm{d}s^n}.$$

- Laplace transform of an integral:

$$\int_0^t f(\tau)\, \mathrm{d}\tau \circ\!\!-\!\!\bullet \frac{F(s)}{s}.$$

---

[1]We use the notation

$$f(0+) := \lim_{\epsilon \to 0} f(\epsilon) \qquad \text{with} \quad \epsilon > 0.$$

- Shift theorems:

$$e^{-at} f(t) \circ\!\!\!-\!\!\!\bullet F(s + a),$$

$$f(t - a) \circ\!\!\!-\!\!\!\bullet e^{-as} F(s) \qquad \text{for} \quad a > 0. \tag{2.26}$$

- Convolution:

$$f * g = g * f \circ\!\!\!-\!\!\!\bullet F(s) \, G(s). \tag{2.27}$$

- Scaling ($a > 0$):

$$f(at) \circ\!\!\!-\!\!\!\bullet \frac{1}{a} \, F\left(\frac{s}{a}\right),$$

$$\frac{1}{a} \, f\left(\frac{t}{a}\right) \circ\!\!\!-\!\!\!\bullet F(as).$$

- Limits:

$$f(0+) = \lim_{s \to \infty} (s \, F(s)),$$

$$f(\infty) := \lim_{t \to \infty} f(t) = \lim_{s \to 0} (s \, F(s)). \tag{2.28}$$

Here f and its derivative must satisfy further requirements [4]. Before using the final-value theorem (2.28), for example, one should verify that the function actually converges for $t \to \infty$.

Like the Fourier transform, the Laplace transform may also be generalized in order to cover distributions (i.e., generalized functions) [1]. Some common Laplace transforms are summarized in Table A.4 on p. 418.

2.3 Transfer Functions

Some dynamical systems[2] may be described by the equation

$$Y(s) = H(s) \, X(s).$$

---

[2]This will be discussed in Sect. 7.1.1.

In this case, $X(s)$ and $Y(s)$ are the Laplace transforms of the input signal $x(t)$ and the output signal $y(t)$, respectively. The Laplace transform $H(s)$ is called the **transfer function** of the system. We discuss two specific input signals:

- Let us assume that the input function $x(t)$ is a **Heaviside step function**

$$\Theta(t) = \begin{cases} 0 & \text{for } t < 0, \\ 1 & \text{for } t \geq 0, \end{cases} \quad \circ\!\!-\!\!\bullet \; \frac{1}{s}.$$

  In this case, the output is

$$Y(s) = \frac{H(s)}{s}.$$

  If we now apply Eq. (2.28), we obtain

$$y(\infty) = \lim_{s \to 0} H(s)$$

  as the long-term **(unit-)step response** of the system.
- If generalized functions are allowed, we may use $x(t) = \delta(t)$ as an input signal. In this case, the correspondence

$$\delta(t) \circ\!\!-\!\!\bullet 1$$

  leads to

$$Y(s) = H(s),$$

  which means that the transfer function $H(s)$ corresponds to the **impulse response** $h(t)$ of the system. The final value of the response $y(t) = h(t)$ is then given by

$$y(\infty) = \lim_{s \to 0} \left(s \, H(s) \right).$$

Let us assume that a system component is specified by the transfer function $H(s)$. If we calculate the **phase response**[3] of this component according to $\varphi(\omega) = \angle H(j\omega)$, the **group delay** can be defined by

$$\tau_{\mathrm{g}} = -\frac{\mathrm{d}\varphi}{\mathrm{d}\omega}.$$

---

[3]The function $A(\omega) = |H(j\omega)|$ is the **amplitude response**. Phase response and amplitude response define the **frequency response** $H(j\omega)$.

Taking a **dead-time element** with $H(s) = e^{-sT_{\text{dead}}}$ (see shift theorem (2.26)) as an example, one obtains the frequency-independent, i.e., constant group delay

$$\tau_{\text{g}} = T_{\text{dead}}.$$

Hence, the dead-time element is an example of a device with linear phase response.

2.4   Mathematical Statistics

The results summarized in this chapter can be read in more detail in [8].

2.4.1   Gaussian Distribution

The **Gaussian distribution** (also called the normal distribution) is given by the **probability density function**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.29}$$

where $\mu, \sigma \in \mathbb{R}$ with $\sigma > 0$ are specified. In order to ensure that $f(x)$ is in fact a valid probability distribution, the equation

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$$

must hold. We show this by substituting

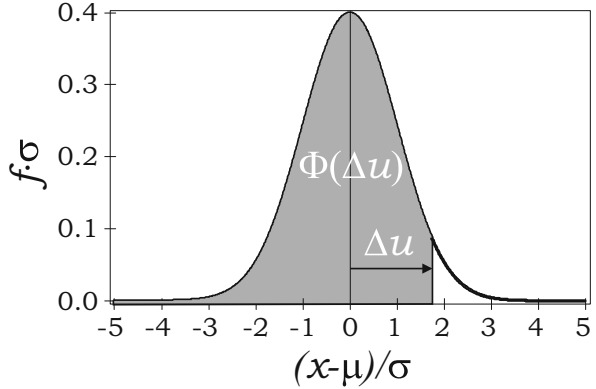$$u = \frac{x - \mu}{\sigma}, \qquad \frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{\sigma}.$$

This leads to

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}u^2}\,\mathrm{d}u.$$

By means of standard methods of mathematical analysis, one may show that

$$\int_{0}^{\infty} e^{-a^2 u^2}\,\mathrm{d}u = \frac{\sqrt{\pi}}{2a},$$

Fig. 2.1 Gaussian
distribution

which actually leads to the result

$$\int_{-\infty}^{+\infty} f(x) \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2} \, du = 1. \tag{2.30}$$

For a given measurement curve that has the shape of a Gaussian distribution, one may use curve-fitting techniques to determine the parameters μ and σ. A simpler method is to determine the **FWHM (full width at half maximum) value**. According to Eq. (2.29), one-half of the maximum value is obtained for

$$e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \overset{!}{=} \frac{1}{2} \qquad \Rightarrow -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 = -\ln 2 \qquad \Rightarrow |x - \mu| = \sigma\sqrt{2 \ln 2}.$$

The FWHM value equals twice this distance (one to the left of the maximum and one to the right of the maximum):

$$\text{FWHM} = \sigma \, 2 \, \sqrt{2 \ln 2} \approx 2.35482 \, \sigma.$$
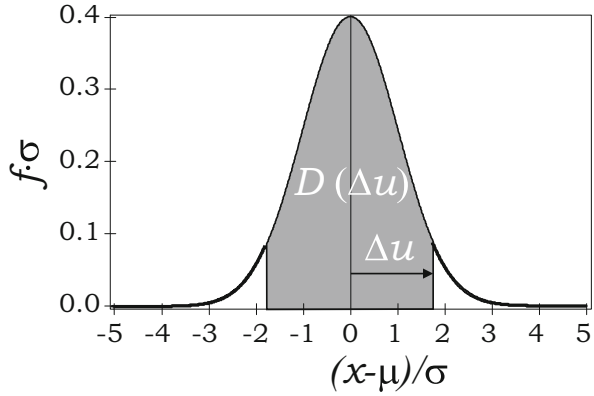
This formula may, of course, lead to less-accurate results than those obtained by the curve-fitting concept if zero line or the maximum cannot be clearly identified in the measurement data.

2.4.2   Probabilities

We now consider the area below the curve $f(x)$ that is located to the left of $x = \mu + \Delta x$, where $\Delta x > 0$ holds. This area will be denoted by Φ:

$$\Phi = \int_{-\infty}^{\mu+\Delta x} f(x) \, dx.$$

Fig. 2.2 Gaussian distribution

It obviously specifies the probability that the **random variable** X is less than $\mu + \Delta x$. By applying the same substitution as that mentioned above, one obtains

$$\Phi = \int_{-\infty}^{\Delta x/\sigma} \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2}u^2} \, du.$$

According to Fig. 2.1 we set

$$\Delta u = \frac{\Delta x}{\sigma}$$

and get

$$\Phi(\Delta u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Delta u} e^{-\frac{1}{2}u^2} \, du.$$

The area D that is enclosed between $\mu - \Delta x$ and $\mu + \Delta x$ (see Fig. 2.2) can be calculated as follows:

$$D(\Delta u) = \Phi(\Delta u) - \Phi(-\Delta u).$$

Due to symmetry, we have

$$\Phi(-\Delta u) = 1 - \Phi(\Delta u),$$

which leads to

$$D(\Delta u) = 2 \, \Phi(\Delta u) - 1.$$

Table 2.2 Integrals of the Gaussian probability density function

Δu	$\Phi_0(\Delta u)$	Δx	$D(\Delta u) = 2\,\Phi_0(\Delta u)$
0	0	0	0
0.5	0.1915	$0.5\,\sigma$	0.3829
1	0.3413	σ	0.6827
1.5	0.4332	$1.5\,\sigma$	0.8664
1.6449	0.45	$1.6449\,\sigma$	0.9000
1.9600	0.475	$1.9600\,\sigma$	0.9500
2	0.4772	$2\,\sigma$	0.9545
2.5	0.4938	$2.5\,\sigma$	0.9876
2.5758	0.4950	$2.5758\,\sigma$	0.9900
3	0.4987	$3\,\sigma$	0.9973
3.2906	0.4995	$3.2906\,\sigma$	0.9990
3.5	0.4998	$3.5\,\sigma$	0.9995

Often, the area Φ_0 is considered, which is located between μ and $\mu + \Delta x$:

$$\Phi_0(\Delta u) = \Phi(\Delta u) - \frac{1}{2}, \qquad \Phi_0(\Delta u) = \frac{1}{\sqrt{2\pi}} \int_0^{\Delta u} e^{-\frac{1}{2}u^2}\, \mathrm{d}u.$$

This shows that D may also be written in the form

$$D(\Delta u) = 2\,\Phi_0(\Delta u).$$

Some examples for these quantities are summarized in Table 2.2.

As an example, the table shows that the random variable is located in the **confidence interval** between $\mu - 2\sigma$ and $\mu + 2\sigma$ with a probability of 95.45%.

2.4.3 Expected Value

Let X be a random variable with probability density function $f(x)$. Then the **expected value** of the function $g(X)$ is given by

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)\, f(x)\, \mathrm{d}x.$$

It is obvious that the expected value is linear:

$$E(a\, g_1(X) + b\, g_2(X)) = a\, E(g_1(X)) + b\, E(g_2(X)).$$

For $g(X) = X^k$, one obtains the kth **moment**:

$$E(X^k) = \int_{-\infty}^{+\infty} x^k\, f(x)\, \mathrm{d}x.$$

By definition, the first moment is the **mean** of the random variable X. For the Gaussian distribution, we obtain

$$E(X) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} x \, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u + \mu) \, e^{-\frac{1}{2}u^2} \, \mathrm{d}u.$$

The term σu in the parentheses leads to an odd integrand, so that this part of the integral vanishes.

Using Eq. (2.30), one obtains the mean

$$E(X) = \mu,$$

which is geometrically obvious.

If we always (not only for the Gaussian distribution) denote the mean by μ, then the kth **central moment** is given by

$$E((X - \mu)^k) = \int_{-\infty}^{+\infty} (x - \mu)^k \, f(x) \, \mathrm{d}x.$$

The second central moment is called the **variance**. For the Gaussian distribution, we obtain

$$E((X-\mu)^2) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-\mu)^2 \, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u)^2 \, e^{-\frac{1}{2}u^2} \, \mathrm{d}u.$$

With

$$a = u \, e^{-u^2/4}, \quad a' = e^{-u^2/4} \left(1 - \frac{u^2}{2}\right),$$

$$b' = u \, e^{-u^2/4}, \quad b = -2 \, e^{-u^2/4},$$

an integration by parts yields

$$\int_{-\infty}^{\infty} u^2 \, e^{-u^2/2} \, \mathrm{d}u = -2u \, e^{-u^2/2} \Big|_{-\infty}^{\infty} + 2 \int_{-\infty}^{+\infty} e^{-u^2/2} \left(1 - \frac{u^2}{2}\right) \, \mathrm{d}u.$$

The first term on the right-hand side vanishes, and we get

$$2 \int_{-\infty}^{\infty} u^2 \, e^{-u^2/2} \, \mathrm{d}u = 2 \int_{-\infty}^{+\infty} e^{-u^2/2} \, \mathrm{d}u.$$

The remaining integral is known from Eq. (2.30):

$$\int_{-\infty}^{\infty} u^2 \, e^{-u^2/2} \, du = \sqrt{2\pi}.$$

Hence we obtain

$$E((X - \mu)^2) = \sigma^2.$$

The variance is generally denoted by σ^2 (not only for the Gaussian distribution), and its square root, the value σ, is called the **standard deviation**.

For a random sample with m values $x_1, x_2, \ldots, x_m$, one defines the **sample mean**

$$\bar{x} = \frac{1}{m} \sum_{k=1}^{m} x_k$$

and the **sample variance**

$$s^2 = \frac{1}{m-1} \sum_{k=1}^{m} (x_k - \bar{x})^2.$$

For large samples, this value does not deviate much from Δx_{rms}^2, where the **root mean square (rms)** is defined[4] as

$$\Delta x_{\text{rms}} = \sqrt{\frac{1}{m} \sum_{k=1}^{m} (x_k - \bar{x})^2}.$$

---

[4]The **root mean square (rms)** of a continuous-time signal $f(t)$ in the interval $[T_1, T_2]$ is defined by

$$f_{\text{rms}} = \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f^2(t) \, dt}. \qquad (2.31)$$

If the time interval is divided into m equal subintervals, one obtains approximately

$$f_{\text{rms}} = \sqrt{\frac{1}{T_2 - T_1} \sum_{k=1}^{m} f_k^2 \frac{T_2 - T_1}{m}}$$

$$\Rightarrow f_{\text{rms}} = \sqrt{\frac{1}{m} \sum_{k=1}^{m} f_k^2} \qquad (2.32)$$

if f_k is regarded as a (time-discrete) sample of $f(t)$ in the kth subinterval. This equation is used in general to define the rms value of a set of values f_k ($k \in \{1, 2, \ldots, m\}$).

2.4.4   Unbiasedness

The individual values x_k of a sample are the observed realizations of the random variables X_k that belong to the same distribution. Also,

$$\bar{X} = \frac{1}{m} \sum_{k=1}^{m} X_k$$

is a random variable for which one may calculate the expected value. From $E(X_k) = \mu$ we obtain

$$E(\bar{X}) = \frac{1}{m} \sum_{k=1}^{m} E(X_k) = \mu,$$

which means that $\bar{X}$ is an **unbiased estimator** of the mean value μ of the population. We now check whether the sample variance

$$S^2 = \frac{1}{m-1} \sum_{k=1}^{m} (X_k - \bar{X})^2$$

is unbiased as well. We have

$$E(S^2) = \frac{1}{m-1} \sum_{k=1}^{m} \left[E(X_k^2) - 2E(X_k \bar{X}) + E(\bar{X}^2) \right]. \qquad (2.33)$$

First of all, we need an expression for $E(X_k^2)$. For this purpose, we point out that all the random variables X_k belong to the same distribution, so that

$$\sigma^2 = E((X_k - \mu)^2) = E(X_k^2) - 2\mu E(X_k) + \mu^2$$

holds. From $E(X_k) = \mu$, we obtain

$$\sigma^2 = E(X_k^2) - \mu^2$$

and

$$E(X_k^2) = \sigma^2 + \mu^2. \qquad (2.34)$$

Now we analyze the second expression in Eq. (2.33), i.e., the expected value of

$$X_k \bar{X} = \frac{1}{m} \sum_{l=1}^{m} X_k X_l.$$

For **independent random variables** X and Y, we have the equation

$$E(XY) = E(X)\,E(Y).$$

In our case, this is satisfied only for $k \neq l$, which means for $m - 1$ terms. The term with $k = l$ leads to the expected value $E(X_k^2)$ derived above. Therefore, we have

$$E(X_k \bar{X}) = \frac{1}{m} \sum_{l=1}^{m} E(X_k X_l) = \frac{1}{m} \left[(m-1)\mu^2 + (\sigma^2 + \mu^2)\right] = \mu^2 + \frac{\sigma^2}{m}.$$

$$(2.35)$$

Finally, we calculate the expected value of

$$\bar{X}^2 = \frac{1}{m^2} \sum_{k=1}^{m} \sum_{l=1}^{m} X_k X_l$$

in an analogous way, obtaining

$$E(\bar{X}^2) = \frac{1}{m^2} \left((m^2 - m)\mu^2 + m(\sigma^2 + \mu^2)\right) = \mu^2 + \frac{\sigma^2}{m}. \qquad (2.36)$$

The results (2.34)–(2.36) may now be used in Eq. (2.33):

$$E(S^2) = \frac{m}{m-1} \left[(\sigma^2 + \mu^2) - 2\left(\mu^2 + \frac{\sigma^2}{m}\right) + \left(\mu^2 + \frac{\sigma^2}{m}\right)\right]$$

$$= \frac{m}{m-1} \left(\sigma^2 - \frac{\sigma^2}{m}\right) = \sigma^2.$$

This shows that the sample variance is an unbiased estimator of the population variance. This is obviously not true for rms values. For large samples, however, this difference is no longer important.

We now calculate the variance of the sample mean $\bar{X}$:

$$E((\bar{X} - \mu)^2) = E(\bar{X}^2) - 2\mu E(\bar{X}) + \mu^2 = \left(\mu^2 + \frac{\sigma^2}{m}\right) - \mu^2 = \frac{\sigma^2}{m}.$$

This shows that an estimate of the population mean from the sample mean becomes better as the sample size becomes larger.

2.4.5   Uniform Distribution

According to

$$f(x) = \begin{cases} \frac{1}{2\Delta x} & \text{for } |x - \mu| \leq \Delta x, \\ 0 & \text{elsewhere,} \end{cases} \qquad \text{with the constant } \Delta x > 0,$$

we now calculate the variance of a uniform distribution:

$$\sigma^2 = E((X - \mu)^2) = \int_{-\infty}^{+\infty} (x - \mu)^2 \, f(x) \, dx = \int_{-\infty}^{+\infty} u^2 \, f(u + \mu) \, du.$$

In the last step, we substituted $u = x - \mu$ to obtain

$$\sigma^2 = \int_{-\Delta x}^{+\Delta x} u^2 \, \frac{1}{2\Delta x} \, du = 2 \int_0^{\Delta x} u^2 \, \frac{1}{2\Delta x} \, du = \frac{1}{\Delta x} \frac{u^3}{3} \Big|_0^{\Delta x} = \frac{\Delta x^2}{3}$$

$$\Rightarrow \sigma = \frac{1}{\sqrt{3}} \, \Delta x.$$

For large samples, we get

$$\Delta x_{\text{rms}} \approx \frac{1}{\sqrt{3}} \, \Delta x.$$

2.5   Bunching Factor

Let us consider a beam signal $I_{\text{beam}}(t)$ of a bunched beam as shown, for example, in Fig. 1.3 on p. 7. The **bunching factor** is defined as

$$\boxed{B_{\text{f}} = \frac{\bar{I}_{\text{beam}}}{I_{\text{beam,max}}},} \qquad (2.37)$$

i.e., it is the ratio of the average beam current to the maximum beam current (cf. Chao [9, Sect. 2.5.3.2, p. 131] or Reiser [10, Sect. 4.5.1, p. 263]). Obviously, the equation

$$\bar{I}_{\text{beam}} = \frac{1}{T_{\text{RF}}} \int_{-T_{\text{RF}}/2}^{T_{\text{RF}}/2} I_{\text{beam}}(t) \, dt$$

holds, where T_{RF} denotes the period.

Now one may replace the true shape of the beam current pulse by a rectangular one with the same maximum value. For $-T_{\mathrm{RF}}/2 < t < T_{\mathrm{RF}}/2$, we then have

$$I_{\mathrm{beam}}(t) = \begin{cases} I_{\mathrm{beam,max}} & \text{for } |t| \leq \tau/2, \\ 0 & \text{elsewhere}, \end{cases}$$

where we have assumed that the bunch is centered at $t = 0$. In this case, one has to choose a pulse width τ in such a way that the same average beam current is obtained:

$$\bar{I}_{\mathrm{beam}} = \frac{\tau}{T_{\mathrm{RF}}}\, I_{\mathrm{beam,max}}.$$

Under these conditions, we obtain the expression

$$B_{\mathrm{f}} = \frac{\tau}{T_{\mathrm{RF}}}$$

for the bunching factor.

We now assume that the beam current pulse has the shape of a Gaussian distribution. This is, of course, possible only if the pulses are significantly shorter than the period time T_{RF}. Under this condition, the beam current will be close to zero before the next pulse starts.

Making use of Eq. (2.29), one may write $I_{\mathrm{beam}}(t)$ in the form

$$I_{\mathrm{beam}}(t) = K\,\frac{1}{\sigma\,\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2} \qquad \text{for} \qquad -T_{\mathrm{RF}}/2 < t < T_{\mathrm{RF}}/2.$$

We have

$$\int_{-\infty}^{+\infty} I_{\mathrm{beam}}(t)\, \mathrm{d}t = K.$$

The average beam current is obtained using the above-mentioned approximation:

$$\bar{I}_{\mathrm{beam}} = \frac{1}{T_{\mathrm{RF}}}\int_{-T_{\mathrm{RF}}/2}^{T_{\mathrm{RF}}/2} I_{\mathrm{beam}}(t)\, \mathrm{d}t \approx \frac{K}{T_{\mathrm{RF}}}.$$

For the maximum current, we obtain

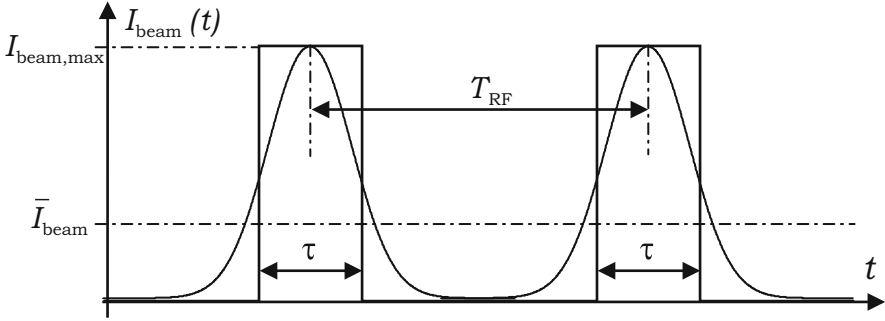$$I_{\mathrm{beam,max}} = \frac{K}{\sigma\,\sqrt{2\pi}},$$

Fig. 2.3 Gaussian beam signal

so that the bunching factor

$$B_f \approx \frac{\sigma \sqrt{2\pi}}{T_{RF}}$$

is obtained. The equivalent length τ of a rectangular pulse is therefore

$$\tau = \sigma \sqrt{2\pi} \approx 2.5 \, \sigma.$$

The two slopes of the rectangular pulse are therefore located at about $\pm 1.25 \, \sigma$. This leads to the conversion between the Gaussian bunch and the rectangular signal that is visualized in Fig. 2.3.

2.6 Electromagnetic Fields

We summarize in this section a few basic formulas that may be found in standard textbooks (cf. [11–17]). We begin with **Maxwell's equations** in their integral form.

In the following, A denotes a two-dimensional domain, and V a three-dimensional domain. For a domain D (two- or threedimensional), ∂D denotes its boundary (with mathematically positive orientation, if applicable).

Maxwell's first equation (**Ampère's law**) in the time domain is

$$\boxed{\oint_{\partial A} \vec{H} \cdot d\vec{r} = \int_A \left(\vec{J} + \dot{\vec{D}} \right) \cdot d\vec{A},} \tag{2.38}$$

where $\vec{H}$ is the magnetizing field, $\vec{J}$ the current density, and $\vec{D}$ the electric displacement field.

Maxwell's second equation in the time domain (**Faraday's law**) reads

$$\oint_{\partial A} \vec{E} \cdot d\vec{r} = -\int_A \dot{\vec{B}} \cdot d\vec{A}. \tag{2.39}$$

Here $\vec{E}$ is the electric field, and $\vec{B}$ is the magnetic field.

Maxwell's third equation states that no magnetic charge exists:

$$\oint_{\partial V} \vec{B} \cdot d\vec{A} = 0. \tag{2.40}$$

The electric charge Q inside a three-dimensional domain V is determined by Maxwell's fourth equation (**Gauss's law**):

$$\oint_{\partial V} \vec{D} \cdot d\vec{A} = \int_V \rho_q \, dV = Q. \tag{2.41}$$

Here, ρ_q denotes the charge density.

The current through a certain region A is given by

$$I = \int_A \vec{J} \cdot d\vec{A},$$

and the voltage along a curve C is defined by

$$V = \int_C \vec{E} \cdot d\vec{r}.$$

Please note that we use the same symbol for voltage and for threedimensional domains, but according to the context this should not lead to confusion.

In material bodies, the simplest relationships (linear isotropic media with relaxation times that are much smaller than the minimum time intervals of interest) between the field vectors are

$$\vec{D} = \epsilon \vec{E}, \qquad \vec{B} = \mu \vec{H}, \qquad \vec{J} = \kappa \vec{E}.$$

The material parameters are the permittivity ϵ, the permeability μ, and the conductivity κ. In vacuum, and approximately also in air, we have

$$\epsilon = \epsilon_0, \qquad \mu = \mu_0.$$

At least for fixed nonmoving domains A, we can write Eq. (2.39) in the form

$$\oint_{\partial A} \vec{E} \cdot d\vec{r} = -\frac{d\Phi_m}{dt},$$

where

$$\Phi_m = \int_A \vec{B} \cdot d\vec{A}$$

is the magnetic flux through the domain A. This form is suitable for induction problems.

Based on the integral form of Maxwell's equations presented above, one may derive their differential form if integral theorems are used:

$$\text{curl } \vec{H} = \vec{J} + \dot{\vec{D}}, \tag{2.42}$$

$$\text{curl } \vec{E} = -\dot{\vec{B}}, \tag{2.43}$$

$$\text{div } \vec{D} = \rho_q, \tag{2.44}$$

$$\text{div } \vec{B} = 0. \tag{2.45}$$

Taking Eq. (2.44) into account, the divergence of Eq. (2.42) leads to the **continuity equation**

$$\text{div}\,\vec{J} + \dot{\rho}_q = 0. \tag{2.46}$$

We will discuss the physical meaning of this equation in Sect. 2.9.

In certain cases (here we assume that domains are filled homogeneously with linear isotropic material), Maxwell's equations may be solved by means of the **vector potential**[5] $\vec{A}$, defined by

$$\vec{B} = \text{curl } \vec{A}, \tag{2.47}$$

and the **scalar potential** Φ, defined by

$$\vec{E} = -\dot{\vec{A}} - \text{grad } \Phi, \tag{2.48}$$

both connected by the **Lorenz gauge condition**

$$\text{div } \vec{A} = -\mu\epsilon\dot{\Phi}. \tag{2.49}$$

---

[5]Please note that we have used the symbol $d\vec{A}$ for the area element and A for two-dimensional domains. The absolute value of the vector potential would be $A = |\vec{A}|$ as well, but the meaning should be clear from the context.

Using these definitions, one obtains the **wave equations**

$$\Delta \vec{A} - \frac{1}{c^2}\ddot{\vec{A}} = -\mu \vec{J},$$ (2.50)

$$\Delta \Phi - \frac{1}{c^2}\ddot{\Phi} = -\frac{\rho_q}{\epsilon}.$$ (2.51)

Here

$$c = \frac{1}{\sqrt{\mu\epsilon}}$$ (2.52)

denotes the speed of light in the material under consideration. The speed of light in vacuum is

$$c_0 = \frac{1}{\sqrt{\mu_0 \epsilon_0}}.$$ (2.53)

For static problems, there is no time dependence of the fields, and according to Maxwell's equations, electric and magnetic fields are therefore decoupled. In this case, the vector potential and the scalar potential also do not depend on time.

Equations (2.48) and (2.51) for homogeneous media thus reduce to

$$\vec{E} = -\text{grad } \Phi$$

and the **Poisson equation**

$$\Delta \Phi = -\frac{\rho_q}{\epsilon},$$ (2.54)

respectively. This equation has to be solved for electrostatic problems.

2.7 Special Relativity

The primary objective of this section is to introduce the nomenclature that is used in this book. This nomenclature is close to that of the introductory text [17] (in German). In any case, the reader should consult standard textbooks on special (and general) relativity (cf. [11, 13, 18–20] in English or [14, 21–28] in German) for an extensive introduction. However, the remainder of the book can also be understood if the formulas presented in this section are regarded as given.

The speed of light c_0 in vacuum has the same value in every **inertial frame**. Therefore, the equation of the wave front

$$x^2 + y^2 + z^2 = c_0^2 t^2$$ (2.55)

in one inertial frame S (e.g., light flash at $t = 0$ at the origin of S) is transformed
into a wave front equation

$$\bar{x}^2 + \bar{y}^2 + \bar{z}^2 = c_0^2 \bar{t}^2$$

that has the same form in a different inertial frame $\bar{S}$. Such transformation behavior
is satisfied by the general **Lorentz transformation**. If one restricts generality in
such a way that at $t = 0$, the origins of the two inertial frames are at the same
position and that one frame $\bar{S}$ moves with constant velocity v in the z-direction
relative to the other frame S, then one obtains the special Lorentz transformation

$$\bar{x} = x, \tag{2.56}$$

$$\bar{y} = y, \tag{2.57}$$

$$\bar{z} = \frac{z - vt}{\sqrt{1 - \frac{v^2}{c_0^2}}}, \tag{2.58}$$

$$\bar{t} = \frac{t - \frac{v}{c_0^2} z}{\sqrt{1 - \frac{v^2}{c_0^2}}}. \tag{2.59}$$

The inverse transformation can be generated if the quantities with a bar (e.g., $\bar{y}$)
are replaced by the same quantities without the bar (e.g., y) and vice versa. In that
case, $\bar{v} = -v$ has to be used (if $\bar{S}$ moves with respect to S with velocity v in the
positive z direction, S will move with respect to $\bar{S}$ in the negative z direction), and c_0
remains the same. This concept for generating inverse transformation formulas may
also be applied to electromagnetic field quantities, whose transformation behavior
is discussed below.

The square root in the denominator of Eqs. (2.58) and (2.59) is typical of
expressions in special relativity. Therefore, the so-called **Lorentz factors** are
defined:

$$\boxed{\beta_v = \frac{v}{c_0},}$$

$$\boxed{\gamma_v = \frac{1}{\sqrt{1 - \beta_v^2}}.}$$

Special relativity may be built up by defining so-called **four-vectors** and **four-
tensors**. For example, the space coordinates are combined with the time "coordi-
nate" in order to define the components of a four-vector that specifies the position
in **space-time**:

$$(\theta^i) = (x, y, z, c_0 t)^{\mathrm{T}} \quad \text{with} \quad i \in \{1, 2, 3, 4\}.$$

Specific values of this four-vector can be interpreted as **events**. In combination with the special choice (**signature**)

$$(g_{ik}) = (g^{ik}) = (\bar{g}_{ik}) = (\bar{g}^{ik}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

for the **metric tensor** ($i, k \in \{1, 2, 3, 4\}$), one obtains the desired transformation behavior of the wave front equation, because

$$\theta^i \theta_i = g_{ik}\theta^i \theta^k = 0,$$

which reproduces Eq. (2.55), is a tensor equation with a tensor of rank 0 (scalar) on the right-hand side. Here we use the **Ricci calculus** and **Einstein's summation convention**. The special Lorentz transformation given above can now be reproduced by

$$\bar{\theta}^i = \bar{a}^i_k \theta^k,$$

which corresponds to the matrix equation

$$(\bar{\theta}^i) = (\bar{a}^i_k) \cdot (\theta^i)$$

if the transformation coefficients

$$(\bar{a}^i_k) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma_v & -\beta_v\gamma_v \\ 0 & 0 & -\beta_v\gamma_v & \gamma_v \end{pmatrix}$$

are chosen (i = row, k = column).

Similarly to the construction of the position four-vector, the vector potential and the scalar potential in electromagnetic field theory may be combined to form the electromagnetic **four-potential** $\mathcal{A}$ according to

$$(\mathcal{A}^i) = (A_x, A_y, A_z, \Phi/c_0)^{\mathrm{T}}.$$

This, for example, allows one to write the Lorenz gauge condition (2.49) for free space in the form

$$\mathcal{A}^i|_i = 0$$

of a tensor equation, where the vertical line indicates a **covariant derivative**, which—in special relativity—corresponds to the partial derivative because the metric coefficients are constant.

The **four-current density** $\mathcal{J}$ is defined by

$$(\mathcal{J}^i) = (J_x, J_y, J_z, \rho_q c_0)^{\mathrm{T}},$$

so that the tensor equation

$$\mathcal{J}^i|_i = 0$$

represents the continuity equation (2.46). The transformation law obviously yields

$$\bar{J}_x = J_x, \tag{2.60}$$

$$\bar{J}_y = J_y, \tag{2.61}$$

$$\bar{J}_z = \gamma_v \left(J_z - v\rho_q \right), \tag{2.62}$$

$$\bar{\rho}_q = \gamma_v \left(\rho_q - \frac{v}{c_0^2} J_z \right). \tag{2.63}$$

With $\vec{v} = v\vec{e}_z$ defining the parallel direction $\parallel$, this may be written in the generalized form

$$\bar{\vec{J}}_\perp = \vec{J}_\perp, \tag{2.64}$$

$$\bar{\vec{J}}_\parallel = \gamma_v \left(\vec{J}_\parallel - \vec{v}\rho_q \right), \tag{2.65}$$

$$\bar{\rho}_q = \gamma_v \left(\rho_q - \frac{\vec{v} \cdot \vec{J}}{c_0^2} \right). \tag{2.66}$$

The **electromagnetic field tensor** may be defined as

$$(\mathcal{B}^{ik}) = \begin{pmatrix} 0 & B_z & -B_y & -E_x/c_0 \\ -B_z & 0 & B_x & -E_y/c_0 \\ B_y & -B_x & 0 & -E_z/c_0 \\ E_x/c_0 & E_y/c_0 & E_z/c_0 & 0 \end{pmatrix},$$

while its counterpart for the other field components in Maxwell's equations may be defined as

$$(\mathcal{H}^{ik}) = \begin{pmatrix} 0 & H_z & -H_y & -c_0 D_x \\ -H_z & 0 & H_x & -c_0 D_y \\ H_y & -H_x & 0 & -c_0 D_z \\ c_0 D_x & c_0 D_y & c_0 D_z & 0 \end{pmatrix},$$

where i specifies the row, and k the column. The introduction of these four-vectors and four-tensors allows one to write Maxwell's equations as[6]

$$\mathcal{H}^{ik}|_i = -\mathcal{J}^k, \tag{2.67}$$

$$\mathcal{B}^{*ik}|_i = 0, \tag{2.68}$$

so that their form remains the same if a Lorentz transformation from one inertial frame to a different one is performed. This form invariance of physical laws is called **covariance**. The covariance of Maxwell's equations implies the constancy of c_0 in different inertial frames, since c_0 is a scalar quantity, a tensor of rank 0. Because $\mathcal{B}^{ik}$ and $\mathcal{H}^{ik}$ are tensors of rank 2, they are transformed according to the transformation rule

$$\bar{\mathcal{B}}^{ik} = \bar{a}_l^i \bar{a}_m^k \mathcal{B}^{lm}, \qquad \bar{\mathcal{H}}^{ik} = \bar{a}_l^i \bar{a}_m^k \mathcal{H}^{lm}.$$

Taking the second transformation rule as an example, this may be translated into the matrix equation

$$(\bar{\mathcal{H}}^{ik}) = (\bar{a}_k^i) \cdot (\mathcal{H}^{ik}) \cdot (\bar{a}_k^i)^{\mathrm{T}}.$$

A long but straightforward calculation then leads to the transformation laws for the corresponding field components:

$$\bar{H}_x = \gamma_v(H_x + vD_y), \tag{2.69}$$

$$\bar{H}_y = \gamma_v(H_y - vD_x), \tag{2.70}$$

$$\bar{H}_z = H_z, \tag{2.71}$$

$$\bar{D}_x = \gamma_v\left(D_x - \frac{\beta_v}{c_0}H_y\right), \tag{2.72}$$

$$\bar{D}_y = \gamma_v\left(D_y + \frac{\beta_v}{c_0}H_x\right), \tag{2.73}$$

$$\bar{D}_z = D_z. \tag{2.74}$$

---

[6]The asterisk is an operator that generates the dual tensor $\mathcal{B}^*$,

$$\mathcal{B}^{*ik} = \frac{1}{2}e^{iklm}\mathcal{B}_{lm},$$

of the tensor $\mathcal{B}$. Here e^{iklm} denotes the complete asymmetric tensor of rank 4.

The generalized form is

$$\vec{\bar{H}}_\perp = \gamma_v \left(\vec{H}_\perp - \vec{v} \times \vec{D}_\perp \right), \tag{2.75}$$

$$\vec{\bar{H}}_\parallel = \vec{H}_\parallel, \tag{2.76}$$

$$\vec{\bar{D}}_\perp = \gamma_v \left(\vec{D}_\perp + \frac{\vec{v} \times \vec{H}_\perp}{c_0^2} \right), \tag{2.77}$$

$$\vec{\bar{D}}_\parallel = \vec{D}_\parallel. \tag{2.78}$$

The remaining transformation laws are obtained analogously:

$$\bar{B}_x = \gamma_v \left(B_x + \frac{v}{c_0^2} E_y \right), \tag{2.79}$$

$$\bar{B}_y = \gamma_v \left(B_y - \frac{v}{c_0^2} E_x \right), \tag{2.80}$$

$$\bar{B}_z = B_z, \tag{2.81}$$

$$\bar{E}_x = \gamma_v (E_x - \beta_v c_0 B_y), \tag{2.82}$$

$$\bar{E}_y = \gamma_v (E_y + \beta_v c_0 B_x), \tag{2.83}$$

$$\bar{E}_z = E_z, \tag{2.84}$$

$$\vec{\bar{B}}_\perp = \gamma_v \left(\vec{B}_\perp - \frac{\vec{v} \times \vec{E}_\perp}{c_0^2} \right), \tag{2.85}$$

$$\vec{\bar{B}}_\parallel = \vec{B}_\parallel, \tag{2.86}$$

$$\vec{\bar{E}}_\perp = \gamma_v \left(\vec{E}_\perp + \vec{v} \times \vec{B}_\perp \right), \tag{2.87}$$

$$\vec{\bar{E}}_\parallel = \vec{E}_\parallel. \tag{2.88}$$

In the scope of this book, there is no need to develop the theory further. Nor do we discuss such standard effects as time dilation, Lorentz contraction, and the transformation of velocities. However, we need some relativistic formulas for mechanics.

The definition of the Lorentz force (1.1) is valid in special relativity—it corresponds to a covariant equation (the charge Q is invariant; it is a scalar quantity).

Also, the equation

$$\boxed{\vec{F} = \frac{d\vec{p}}{dt}}$$

with the momentum definition

$$\vec{p} = m\,\vec{u}$$

based on the velocity

$$\vec{u} = \frac{d\vec{r}}{dt}$$

still holds. However, the mass m is not invariant. Only the **rest mass** m_0 is a tensor of rank zero, i.e., a scalar:

$$m = \frac{m_0}{\sqrt{1 - \frac{u^2}{c_0^2}}} = m_0\gamma_u.$$

Please note that we strictly distinguish between the velocities v and u and also between the related Lorentz factors. The velocity $\vec{v} = v\,\vec{e}_z$ is defined as the relative velocity of the inertial frame $\bar{S}$ with respect to the inertial frame S, i.e., the velocity between these two reference frames. The velocity $\vec{u}$ is the velocity of a particle measured in the first inertial frame S. Consequently, $\bar{\vec{u}}$ is the velocity of the same particle in the inertial frame $\bar{S}$. If it is clear what is meant by a certain Lorentz factor, one may, of course, omit the subscript.

The **total energy** of a particle with velocity $\vec{u}$ is given by

$$W_{\text{tot}} = m\,c_0^2 = \gamma_u m_0 c_0^2. \tag{2.89}$$

Consequently, the **rest energy** is obtained for $\vec{u} = 0$, which leads to $\gamma_u = 1$:

$$W_{\text{rest}} = m_0 c_0^2.$$

Therefore, the **kinetic energy** is

$$W_{\text{kin}} = W_{\text{tot}} - W_{\text{rest}} = m_0 c_0^2(\gamma_u - 1).$$

Using the Lorentz factors, one may write the momentum in the form

$$\vec{p} = m\vec{u} = m_0 c_0 \vec{\beta}_u \gamma_u,$$

leading to the absolute value

$$p = mu = m_0 c_0 \beta_u \gamma_u. \tag{2.90}$$

Here we used the definition

$$\vec{\beta}_u = \frac{\vec{u}}{c_0}. \tag{2.91}$$

If we have a look at Eqs. (2.89)–(2.91), we observe that γ is related to the energy, the product $\beta\gamma$ to the momentum, and β corresponds to the velocity. It is often helpful to keep this correspondence in mind when complicated expressions containing a large number of Lorentz factors are evaluated. One should also keep in mind that when one of the expressions β, γ, $\beta\gamma$ is known, the others are automatically fixed as well.

This is why we can also convert expressions for relative deviations into each other. For example, we may calculate the time derivative of

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{2.92}$$

as follows:

$$\dot{\gamma} = -\frac{-2\beta\dot{\beta}}{2\left(1-\beta^2\right)^{3/2}} = \beta\gamma^3\dot{\beta}$$

$$\Rightarrow \frac{\dot{\gamma}}{\gamma} = \beta^2\gamma^2\frac{\dot{\beta}}{\beta}.$$

Here we can use the relation

$$\gamma^2 - \beta^2\gamma^2 = 1, \tag{2.93}$$

which follows directly from Eq. (2.92):

$$\frac{\dot{\gamma}}{\gamma} = (\gamma^2 - 1)\frac{\dot{\beta}}{\beta}.$$

Expressions of this type are very helpful, because they can be translated as follows:

$$\frac{\Delta W_{\text{tot}}}{W_{\text{tot}}} = (\gamma^2 - 1)\frac{\Delta u}{u}.$$

This conversion is possible if the relative change in the quantities is sufficiently small. In the example presented here, one can see directly that a velocity deviation of 1% is transformed into an energy deviation of 3% if $\gamma = 2$ holds.

As a second example, we can calculate the time derivative of Eq. (2.93):

$$2\gamma\dot{\gamma} - 2(\beta\gamma)\frac{d(\beta\gamma)}{dt} = 0$$

$$\Rightarrow \frac{\dot{\gamma}}{\gamma} = \beta^2\frac{1}{(\beta\gamma)}\frac{d(\beta\gamma)}{dt}.$$

Table 2.3 Conversion of relative deviations

$\dfrac{\Delta\beta}{\beta}$	$= \dfrac{\Delta\beta}{\beta}$	$= \dfrac{1}{\gamma^2-1}\dfrac{\Delta\gamma}{\gamma} = \dfrac{1}{(\beta\gamma)^2}\dfrac{\Delta\gamma}{\gamma}$	$= \gamma^{-2}\dfrac{\Delta(\beta\gamma)}{\beta\gamma}$	$= \dfrac{1}{\gamma(\gamma+1)}\dfrac{\Delta\gamma}{\gamma-1}$
$\dfrac{\Delta\gamma}{\gamma}$	$= (\gamma^2-1)\dfrac{\Delta\beta}{\beta} = (\beta\gamma)^2\dfrac{\Delta\beta}{\beta}$	$= \dfrac{\Delta\gamma}{\gamma}$	$= \beta^2\dfrac{\Delta(\beta\gamma)}{\beta\gamma}$	$\left(1-\gamma^{-1}\right)\dfrac{\Delta\gamma}{\gamma-1}$
$\dfrac{\Delta(\beta\gamma)}{\beta\gamma}$	$= \gamma^2\dfrac{\Delta\beta}{\beta}$	$= \beta^{-2}\dfrac{\Delta\gamma}{\gamma}$	$= \dfrac{\Delta(\beta\gamma)}{\beta\gamma}$	$= \dfrac{\gamma}{\gamma+1}\dfrac{\Delta\gamma}{\gamma-1}$
$\dfrac{\Delta\gamma}{\gamma-1}$	$= \gamma(\gamma+1)\dfrac{\Delta\beta}{\beta}$	$= \dfrac{\gamma}{\gamma-1}\dfrac{\Delta\gamma}{\gamma}$	$= \left(1+\gamma^{-1}\right)\dfrac{\Delta(\beta\gamma)}{\beta\gamma}$	$= \dfrac{\Delta\gamma}{\gamma-1}$

Note: $\dfrac{\Delta W_{\text{tot}}}{W_{\text{tot}}} = \dfrac{\Delta\gamma}{\gamma}$, $\quad \dfrac{\Delta p}{p} = \dfrac{\Delta(\beta\gamma)}{\beta\gamma}$, $\quad \dfrac{\Delta W_{\text{kin}}}{W_{\text{kin}}} = \dfrac{\Delta\gamma}{\gamma-1}$

This can be translated into

$$\frac{\Delta W_{\text{tot}}}{W_{\text{tot}}} = \beta^2 \frac{\Delta p}{p}.$$

Relations like these are summarized in Table 2.3.

In accelerator physics and engineering, specific units that contain the elementary charge e are often used to specify the energy of the beam. This is due to the fact that the energy that is gained by a charge[7] $Q = z_q e$ is given by formula (1.2),

$$\Delta W = QV = z_q eV.$$

An electron that passes a voltage of $V = 1\,\text{kV}$ will therefore lose or gain an energy of $1\,\text{keV}$, depending on the orientation of the voltage. We have only to insert the quantities into the formula without converting e into SI units. In order to convert an energy that is given in eV into SI units, one simply has to insert $e = 1.6022\cdot10^{-19}\,\text{C}$, so that $1\,\text{eV} = 1.6022 \cdot 10^{-19}\,\text{J}$ holds.

Also, the rest energy of particles is often specified in eV. For example, the electron rest mass $m_{\text{e}} = 9.1094\cdot10^{-31}\,\text{kg}$ corresponds to an energy of $510.999\,\text{keV}$.

As we saw above, the energy directly determines the Lorentz factors and the velocity. Therefore, it is desirable to specify the energy in a unit that directly corresponds to a certain velocity. Due to

$$W_{\text{kin}} = mc_0^2 - m_0 c_0^2 = m_0 c_0^2(\gamma - 1),$$

a kinetic energy of $1\,\text{MeV}$ leads to different values for γ if different particle rest masses m_0 are considered. This is why one introduces another energy unit for ions. An ion with mass number A has rest mass

$$m_0 = A_{\text{r}} m_{\text{u}},$$

---

[7]For ions, z_q is the **charge number**. For electrons, one has to set $z_q = -1$, while for protons and positrons, $z_q = 1$ has to be defined.

where $m_u = 1.66054 \cdot 10^{-27}$ kg denotes the **unified atomic mass unit** (as mentioned below, A_r differs slightly from A). Therefore, one obtains

$$W_{\text{kin,u}} = \frac{W_{\text{kin}}}{A_r} = m_u c_0^2 (\gamma - 1).$$

If the value on the right-hand side is specified now, γ is determined in a unique way, since m_u and c_0 are global constants. As an example, an ion beam with a kinetic energy of[8] 11.4 MeV/u corresponds to $\gamma = 1.0122386$ and $\beta = 0.15503$. We do not need to specify the ion species.

Ions are usually specified by the notation

$$^A_Z\text{Element}^{z_q+}.$$
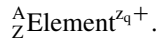
Here A is the (integer) **mass number**, i.e., the number of **nucleons** (protons plus neutrons); Z is the **atomic number**, which equals the number of protons and identifies the element. For example,

$$^{238}_{92}\text{U}^{28+}$$

indicates a uranium ion that has $A - Z = 146$ neutrons. Different uranium **isotopes**[9] exist with a different number of neutrons. The number of protons however, is the same for all these isotopes. Therefore, Z is redundant information that is already included in the element name. In the last example, the uranium atom has obviously lost 28 of its 92 electrons, leading to the charge number $z_q = 28$.

The unified atomic mass unit m_u is defined as $1/12$ of the mass of the atomic nucleus $^{12}_6$C. For different ion species and isotopes, the mass is not exactly an integer multiple of m_u (reasons: different mass of protons and neutrons, relativistic mass defect due to binding energy). For ^{238}U, for example, one has $A_r = 238.050786$, which approximately equals $A = 238$.

2.8 Nonlinear Dynamics

A continuous **dynamical system** of first order may be described by the following first-order **ordinary differential equation (ODE)**:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x, t).$$

---

[8] read: MeV per nucleon.

[9] A **nuclide** is specified by the number of protons Z and the number of neutrons $A - Z$. The chemical element is determined by the number of protons only. Different nuclides belonging to the same chemical element are called isotopes (of that element).

The state of a dynamical system of order n is represented by the values of n variables $x_1, x_2, \ldots, x_n$, which may be combined into a vector $\vec{r} = (x_1, x_2, \ldots, x_n)$. Hence, a dynamical system of order n is described by the system of ordinary differential equations

$$\boxed{\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}, t).} \qquad (2.94)$$

One should note that the system of ODEs is still of order 1, but of dimension n. Such a system is called **autonomous** when $\vec{v}(\vec{r}, t)$ does not depend on the time[10] t, i.e., when

$$\vec{v}(\vec{r}, t) = \vec{v}(\vec{r})$$

holds. The next sections will show that Eq. (2.94), which may look very simple at first sight, includes a huge variety of problems.

2.8.1 Equivalence of Differential Equations and Systems of Differential Equations

Let us consider the nth-order linear ordinary differential equation

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}(t)\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1(t)\frac{\mathrm{d}y}{\mathrm{d}t} + a_0(t)y(t) = b(t)$$

with dimension 1. One sees that by means of the definitions

$$x_1 = y,$$
$$x_2 = \frac{\mathrm{d}y}{\mathrm{d}t},$$
$$\cdots$$
$$x_n = \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}},$$

it may be converted into the form

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = x_3,$$
$$\cdots$$
$$\dot{x}_n = b(t) - a_0(t)x_1 - a_1(t)x_2 - \cdots - a_{n-1}(t)x_n,$$

---

[10]The variable t is not necessarily the time, but we will often call it the time in order to make the interpretation easier.

which is equivalent to the standard form

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}, t).$$

If b and all the a_k do not depend on time (ODE of order n with constant coefficients), then $\vec{v}$ will also not depend on time explicitly, so that an autonomous system is present.

Although the vector field $\vec{v}$ is called a velocity function, it does not always correspond to a physical velocity. As already mentioned, the variable t is not necessarily the physical time. However, we will use this notation because the reader may always interpret these variables in terms of the mechanical analogy, which may help to understand the physical background.

The above-mentioned equivalence is also valid for nonlinear ODEs of the form

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} = F\left(t, y, \frac{\mathrm{d}y}{\mathrm{d}t}, \dots, \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}}\right),$$

where[11] $F \in C^1$. Also here, we may use

$$x_1 = y,$$
$$x_2 = \frac{\mathrm{d}y}{\mathrm{d}t},$$
$$\dots$$
$$x_n = \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}},$$

to obtain the standard form

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = x_3,$$
$$\dots$$
$$\dot{x}_n = F(t, x_1, x_2, \dots, x_n)$$

$$\Leftrightarrow \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}, t).$$

---

[11]In this book, C^k denotes the class of functions that are k-times **continuously differentiable**: C^0 is the class of **continuous** functions, C^1 the class of continuously differentiable functions.

An autonomous system results when F and $\vec{v}$ do not explicitly depend on the time variable t.

2.8.2   Autonomous Systems

Hereinafter, we will consider only autonomous systems if a time dependence is not stated explicitly.

2.8.2.1   Time Shift

An advantage of autonomous systems is the fact that if a solution $y(t)$ of

$$\frac{d^n y}{dt^n} = F\left(y, \frac{dy}{dt}, \ldots, \frac{d^{n-1} y}{dt^{n-1}}\right)$$

is known, then $z(t) = y(t - T)$ will also be a solution if T is a constant time shift. This can be shown as follows:

The solution $y(t)$ is the first component of the vector $\vec{r}(t)$ that satisfies the differential equation[12]

$$\frac{d\vec{r}}{dt} = \vec{v}(\vec{r}).$$

Therefore, $z(t)$ is the first component of the vector

$$\vec{r}_{\text{shift}}(t) = \vec{r}(t - T).$$

We obtain

$$\frac{d\vec{r}_{\text{shift}}}{dt} = \frac{d\vec{r}}{dt}\bigg|_{t-T} = \vec{v}(\vec{r})\big|_{t-T} = \vec{v}(\vec{r}_{\text{shift}}).$$

One sees that $\vec{r}_{\text{shift}}(t)$ satisfies the system of ODEs in the same way as $\vec{r}(t)$ does. Due to the equivalence with the differential equation of order n, $z(t)$ will be a solution as well.

This explains, for instance, why $\sin(\omega t)$ must be a solution of the homogeneous differential equation

$$\ddot{y} + \omega^2 y = 0$$

---

[12]In the text, we will not explicitly distinguish between systems of ordinary differential equations and ordinary differential equations, since this difference should be obvious based on the notation.

if one knows that $\cos(\omega t)$ is a solution. This ODE is autonomous, and these two solutions differ only by a time shift.

2.8.2.2   Phase Space

The **phase space** may be defined as the continuous space of all possible states of a dynamical system. In our case, the dynamical system is described by an autonomous system of ordinary differential equations.

The graphs of the solutions $\vec{r}(t)$ of the differential equation are the **integral curves** or **solution curves** in the n-dimensional phase space. Such an integral curve contains the dependence on the parameter t (which is usually but not necessarily the time). A different parameterization therefore leads to a different integral curve.

The set of all image points of the map $t \mapsto \vec{r}(t)$ is called the **orbit**. An orbit does not contain dependence on the parameter t. A different parameterization therefore leads to the same orbit, since the same image points are obtained simply by a different value of the parameter t.

Different orbits of an autonomous system are often drawn in a **phase portrait**, which may be defined as the set of all orbits.

2.8.3   Existence and Uniqueness of the Solution of Initial Value Problems

The standard form

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r})$$

has the advantage that it can be solved numerically according to the (explicit) **Euler method**:

$$\vec{r}_{k+1} = \vec{r}_k + \Delta t \cdot \vec{v}(\vec{r}_k)$$

$$t_k = t_0 + k \; \Delta t.$$

It is obvious that by defining the initial condition

$$\vec{r}_0 = \vec{r}(t_0),$$

the states

$$\vec{r}_k \approx \vec{r}(t_k) = \vec{r}(t_0 + k\Delta t)$$

of the system at different times can be derived iteratively for $k > 0$ ($k \in \mathbb{N}$). The states of the system may be calculated for both future times $t > t_0$ and past times $t < t_0$ in a unique way by selecting the sign of Δt. However, this is possible only in a certain neighborhood around t_0, as we will see in the next sections.

It is obvious that by defining $\vec{r}_0$, n scalar initial conditions are required to make the solution unique.

2.8.3.1   Existence of a Local Solution

The existence of a solution is ensured by the following theorem:

Theorem 2.1 (Peano).   *Consider an initial value problem*

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}, t) \qquad \vec{r}(t_0) = \vec{r}_0$$

with a continuous $\vec{v} : D \rightarrow \mathbb{R}^n$ on an open set $D \subset \mathbb{R}^{n+1}$. Then there exists $\alpha(\vec{r}_0, t_0) > 0$ such that the initial value problem has at least one solution in the interval $[t_0 - \alpha, t_0 + \alpha]$.

(See Aulbach [29, Theorem 2.2.3].)

Remark. We may easily see that $\vec{v}$ must be continuous. If we choose $v = \Theta(t)$ (Heaviside step function) in the one-dimensional case, we immediately see that the derivative $\frac{\mathrm{d}r}{\mathrm{d}t}$ is not defined at $t = 0$. Therefore, in the scope of classical analysis, we have to exclude functions that are not continuous. In the scope of distribution theory, the solution $r = t \,\Theta(t)$ is obvious.

2.8.3.2   Uniqueness of a Local Solution

Uniqueness can be ensured if the vector field $\vec{v}$ satisfies a **Lipschitz condition** or if it is continuously differentiable.

Definition 2.2.   The vector function $\vec{v}(\vec{r}, t) : D \rightarrow \mathbb{R}^n$ ($D \subset \mathbb{R}^{n+1}$ open) is said to satisfy a global Lipschitz condition on D with respect to $\vec{r}$ if there is a constant $K > 0$ such that for all $(\vec{r}_1, t), (\vec{r}_2, t) \in D$, the condition

$$\left\| \vec{v}(\vec{r}_1, t) - \vec{v}(\vec{r}_2, t) \right\| \leq K \left\| \vec{r}_1 - \vec{r}_2 \right\|$$

holds. Instead of saying that a function satisfies a global Lipschitz condition, one also speaks of a function that is Lipschitz continuous.

(Cf. Aulbach [29, Definition 2.3.5] and Perko [30, p. 71, Definition 2].)

Definition 2.3.   The vector function $\vec{v}(\vec{r}, t) : D \rightarrow \mathbb{R}^n$ ($D \subset \mathbb{R}^{n+1}$ open) is said to satisfy a local Lipschitz condition on D with respect to $\vec{r}$ if for each $(\vec{r}_0, t_0) \in D$,

there exist a neighborhood $U_{(\vec{r}_0,t_0)} \subset D$ of $(\vec{r}_0, t_0)$ and a constant $K > 0$ such that for all $(\vec{r}_1, t), (\vec{r}_2, t) \in U_{(\vec{r}_0,t_0)}$, the condition

$$\left\| \vec{v}(\vec{r}_1, t) - \vec{v}(\vec{r}_2, t) \right\| \leq K \left\| \vec{r}_1 - \vec{r}_2 \right\|$$

holds. Instead of saying that a function satisfies a local Lipschitz condition, one also speaks of a function that is locally Lipschitz continuous.

(Cf. Aulbach [29, Definition 2.3.5], Wirsching [31, Definition 3.4], and Perko [30, p. 71, Definition 2].)

In other words, the function satisfies a local Lipschitz condition if for every point, we can find a neighborhood such that a "global" Lipschitz condition holds in that neighborhood.

Example. The function $f(x) = x^2$ is locally Lipschitz continuous, but it is not Lipschitz continuous.

Theorem 2.4 (Picard–Lindelöf). *Consider the initial value problem*

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}, t) \qquad \vec{r}(t_0) = \vec{r}_0$$

with continuous $\vec{v} : D \to \mathbb{R}^n$ ($D \subset \mathbb{R}^{n+1}$ open). Suppose that the vector function $\vec{v}(\vec{r}, t)$ is locally Lipschitz continuous with respect to $\vec{r}$. Then there exists $\alpha(\vec{r}_0, t_0) > 0$ such that the initial value problem has a unique solution in the interval $[t_0 - \alpha, t_0 + \alpha]$.

(See Aulbach [29, Theorem 2.3.7].)

Every locally Lipschitz continuous function is also continuous.

Every continuously differentiable function satisfies a local Lipschitz condition, i.e., is locally Lipschitz continuous (Aulbach [29, p. 77], Arnold [32, p. 279], Perko [30, lemma on p. 71]).

Therefore, the Picard–Lindelöf theorem may simply be rewritten for continuously differentiable functions instead of locally Lipschitz continuous functions (Perko [30, p. 74]: "The Fundamental Existence-Uniqueness Theorem," Guckenheimer/Holmes [33, Theorem 1.0.1]).

2.8.3.3 Maximal Interval of Existence

One may try to make the solution interval larger by using the endpoint of the solution interval as a new initial condition. If this strategy is executed iteratively, one obtains the **maximal interval of existence**. It is an open interval (cf. [30, p. 89, Theorem 1]). The maximal interval of existence does not necessarily correspond to the full real time axis. Further requirements are necessary to ensure this.

2.8.3.4 Global Solution

A continuously differentiable vector field $\vec{v}$ is called **complete** if it induces a **global flow**,[13] i.e., if its integral curves are defined for all $t \in \mathbb{R}$.

Every differentiable vector field with compact support is complete.

The following theorem shows that certain restrictions on the "velocity" $\vec{v}(\vec{r})$ are sufficient for completeness:

Theorem 2.5. *Let the vector function $\vec{v}(\vec{r})$ with $\vec{v} : D \rightarrow \mathbb{R}^n$ ($D \subset \mathbb{R}^n$ open) be continuously differentiable and linearly bounded with $K, L \geq 0$:*

$$\left\| \vec{v}(\vec{r}) \right\| \leq K \left\| \vec{r} \right\| + L.$$

Then the initial value problem

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}) \qquad \vec{r}(t_0) = \vec{r}_0$$

has a global flow.

(Cf. Zehnder [34, Proposition IV.3, p. 130], special form of Theorem 2.5.6, Aulbach [29].)

According to Amann [35, Theorem 7.8], the solution will then be bounded for finite time intervals.

Like many other authors, Perko [30, p. 188, Theorem 3] requires that $\vec{v}(\vec{r})$ satisfy a global Lipschitz condition

$$\left\| \vec{v}(\vec{r}_1) - \vec{v}(\vec{r}_2) \right\| \leq K \left\| \vec{r}_1 - \vec{r}_2 \right\|$$

for arbitrary $\vec{r}_1, \vec{r}_2 \in \mathbb{R}^n$. For $\vec{r}_2 = 0$, this leads to linear boundedness, as one may show by means of the reverse triangle inequality, but it is a stronger condition.

Example. The ODE

$$\dot{y} = 1 + y^2$$

is obviously satisfied for

$$y = \tan\ t = \frac{\sin t}{\cos t} \qquad \dot{y} = \frac{\cos^2 t + \sin^2 t}{\cos^2 t} = 1 + \tan^2 t.$$

This solution may be found by separation of variables. An arbitrary initial condition $y(0) = y_0$ may be satisfied if the shifted solution

$$y = \tan(t + \tau)$$

---

[13]The exact definition of a flow will be given in Sect. 2.8.6 on p. 61.

is considered. In any case, however, the solution curve reaches infinity while t is still finite. The "vector" field $v(y) = 1 + y^2$ is not complete, and it is obviously not linearly bounded.

If we simplify the results of this section, we may summarize them as follows:

- The existence of a local solution is ensured by continuity of $\vec{v}$.
- Local Lipschitz continuity ensures uniqueness of the solution. If $\vec{v}$ is continuously differentiable, uniqueness is also guaranteed.
- If linear boundedness of $\vec{v}$ is required in addition, a global solution/global flow exists.

For the sake of simplicity, we will consider only complete vector fields in the following.

2.8.3.5   Linear Systems of Ordinary Differential Equations

For linear systems of differential equations with

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = A \cdot \vec{r}, \qquad \vec{r}(t_0) = \vec{r}_0,$$

where A is a quadratic matrix with real constant elements, we may use the **matrix norm**:

$$\left\| \vec{v}(\vec{r}) \right\| = \left\| A \cdot \vec{r} \right\| \leq \|A\| \left\| \vec{r} \right\| = K \left\| \vec{r} \right\|.$$

Therefore, the conditions of Theorem 2.5 are satisfied, and a unique solution with a global flow exists. One may specifically use the **Frobenius norm**

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{n} \sum_{k=1}^{n} |a_{ik}|^2},$$

which is compatible with the **Euclidean norm**

$$\left\| \vec{r} \right\| = \sqrt{\sum_{i=1}^{n} |r_i|^2}$$

of a vector, so that

$$\left\| A \cdot \vec{r} \right\| \leq \|A\|_{\mathrm{F}} \left\| \vec{r} \right\|$$

holds.

2.8.4   Orbits

Two distinct orbits of an autonomous system do not intersect. In order to prove this, we assume the contrary. Suppose that two distinct orbits defined by $\vec{r}_1(t)$ and $\vec{r}_2(t)$ intersect according to

$$\vec{r}_1(t_1) = \vec{r}_2(t_2).$$

Please note that the intersection point may be reached for different values t_1 and t_2 of the parameter t, since we require only that the orbits (i.e., the images of the solution curves) intersect. As shown in Sect. 2.8.2.1,

$$\vec{r}_{\text{shift}}(t) = \vec{r}_1(t + t_1 - t_2)$$

is also a solution of the differential equation. Therefore, we have

$$\vec{r}_{\text{shift}}(t_2) = \vec{r}_1(t_1) = \vec{r}_2(t_2).$$

Hence $\vec{r}_{\text{shift}}(t)$ and $\vec{r}_2(t)$ satisfy the same initial conditions at the time t_2. This means that the solution curves $\vec{r}_{\text{shift}}(t)$ and $\vec{r}_2(t)$ are identical.

Since $\vec{r}_1(t)$ is simply time-shifted with respect to $\vec{r}_{\text{shift}}(t) = \vec{r}_2(t)$, the images, i.e., the orbits, will be identical. This means that two orbits are completely equal if they have one point in common.

In other words, each point of phase space is crossed by only one orbit.

2.8.5   Fixed Points and Stability

Vectors $\vec{r} = \vec{r}_{\text{F}}$ for which

$$\vec{v}(\vec{r}) = 0$$

holds are called **fixed points** (or **equilibrium points** or **stationary points** or **critical points**) of the dynamical system given by

$$\frac{d\vec{r}}{dt} = \vec{v}(\vec{r}).$$

This nomenclature is obvious, since a particle that is initially located at

$$\vec{r}(t_0) = \vec{r}_{\text{F}}$$
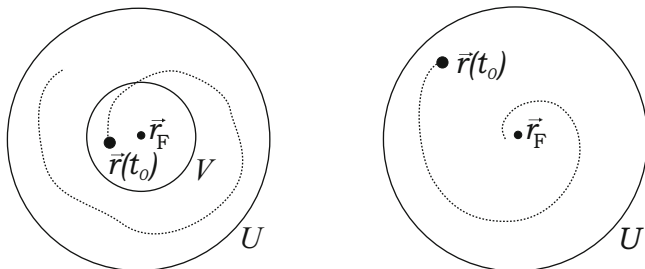
Fig. 2.4 Stability (*left*) and asymptotic stability (*right*) of a fixed point

will stay there forever:

$$\vec{r}(t) = \vec{r}_{\mathrm{F}} \qquad \text{for } t > t_0.$$

Definition 2.6. A fixed point of an autonomous dynamical system is called an **isolated fixed point** or a **nondegenerate fixed point** if an environment of the fixed point exists that does not contain any other fixed points.

(Cf. Sastry [36, Definition 1.4, p. 13], Perko [30, Definition 2, p. 173].)

We now define the stability of fixed points according to Lyapunov.

Definition 2.7. A fixed point is called **stable** if for every neighborhood U of $\vec{r}_{\mathrm{F}}$, another neighborhood $V \subset U$ of $\vec{r}_{\mathrm{F}}$ exists such that a trajectory starting in V at $t = t_0$ will remain in U for all $t \geq t_0$ (see Fig. 2.4). Otherwise, the fixed point is called **unstable**.

Please note that it is usually necessary to choose V smaller than U, because the shape of the orbit may cause the trajectory to leave U for some starting points in U even if $\vec{r}_{\mathrm{F}}$ is stable.

Definition 2.8. A stable fixed point $\vec{r}_{\mathrm{F}}$ is called **asymptotically stable** if a neighborhood U of $\vec{r}_{\mathrm{F}}$ exists such that for every trajectory that starts at $t = t_0$ in U, the following equation holds:

$$\lim_{t \to \infty} \vec{r}(t) = \vec{r}_{\mathrm{F}}.$$

(See, e.g., Perko [30, Definition 1, p. 129].)

Definition 2.9. A function $L(\vec{r})$ with $L \in C^1$ and $L : U \to \mathbb{R}$ ($U \subset \mathbb{R}^n$ open) is called a **Lyapunov function** for the fixed point $\vec{r}_{\mathrm{F}}$ of the autonomous system

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}) \qquad \vec{v} \in C^1(D) \qquad D \subset \mathbb{R}^n \text{open}$$

if

$$L(\vec{r}_{\mathrm{F}}) = 0$$

and

$$L(\vec{r}) > 0 \qquad \text{for} \quad \vec{r} \in U \setminus \{\vec{r}_{\mathrm{F}}\},$$

$$\dot{L} = \vec{v} \cdot \operatorname{grad} L \leq 0 \qquad \text{for} \quad \vec{r} \in U \setminus \{\vec{r}_{\mathrm{F}}, \}$$

hold in a neighborhood $U \subset D$ of $\vec{r}_{\mathrm{F}}$.

A Lyapunov function is called a **strict Lyapunov function** if

$$\dot{L} = \vec{v} \cdot \operatorname{grad} L < 0 \qquad \text{for} \quad \vec{r} \in U \setminus \{\vec{r}_{\mathrm{F}}\}$$

holds.

(Cf. Perko [30, p. 131, Theorem 3], La Salle/Lefschetz [37, Sect. 8], Gucken-heimer/Homes [33, Theorem 1.0.2].)

Theorem 2.10. *If a Lyapunov function for a fixed point $\vec{r}_{\mathrm{F}}$ of an autonomous system exists, then this fixed point $\vec{r}_{\mathrm{F}}$ is stable. If a strict Lyapunov function exists, then this fixed point $\vec{r}_{\mathrm{F}}$ is asymptotically stable.*

(Cf. Perko [30, p. 131, Theorem 3].)

It is easy to see that this theorem is valid. For two-dimensional systems with the particle trajectory $\vec{r}(t) = x(t)\,\vec{e}_x + y(t)\,\vec{e}_y$, we obtain, for example,[14]

$$\dot{L} = \frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial L}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = \vec{v} \cdot \operatorname{grad} L.$$

If this expression is negative, the strict Lyapunov function will decrease while the particle continues on its path. Since the minimum of the Lyapunov function is obtained for $\vec{r}_{\mathrm{F}}$, it is clear that the particle will move toward the fixed point.

Similar reasoning applies for a Lyapunov function that is not strict. In this case, the particle cannot move away from the fixed point, because the Lyapunov function does not increase. However, it will not necessarily get closer to the fixed point.

---

[14] According to the definition, the Lyapunov function depends on the location (x, y). The chain rule that is applied here implies that we are following the trajectory $\vec{r}(t)$, so that $\dot{L}$ depends only on the time t.

The gradient of L points in the direction in which L increases. Therefore, the scalar product $\vec{v} \cdot \operatorname{grad} L$ becomes negative if $\vec{v}$ has a component in the opposite direction (decreasing L).

2.8.6   *Flows of Linear Autonomous Systems*

Having shown above that a linear autonomous system possesses a global flow, we shall now compute this flow. If an autonomous system of order n is linear, we may describe it by

$$\frac{d\vec{r}}{dt} = \vec{v}(\vec{r})$$

with

$$\vec{v}(\vec{r}) = A \cdot \vec{r},$$

where A is a quadratic $n \times n$ matrix with real constant elements. The ansatz

$$\vec{r} = \vec{w}\, e^{\lambda t}$$

leads to

$$\lambda \vec{w} = A \cdot \vec{w},$$

or

$$(A - \lambda I) \cdot \vec{w} = 0.$$

For nontrivial solutions $\vec{w} \neq 0$, the condition

$$\boxed{\det(A - \lambda I) = 0}$$

is necessary, which determines the **eigenvalues** λ. For the sake of simplicity, we now assume that all n eigenvalues are distinct and that there is one **eigenvector** belonging to each eigenvalue (A is diagonalizable in this case). The overall solution of the homogeneous system of ODEs may then be written in the form

$$\vec{r}(t) = \sum_{k=1}^{n} C_k\, \vec{w}_k e^{\lambda_k t}, \tag{2.95}$$

where $\vec{w}_k$ denotes the eigenvector that belongs to the eigenvalue $\lambda = \lambda_k$ and where the C_k are constants. For the initial condition at $t = 0$, we therefore have

$$\vec{r}_0 = \vec{r}(0) = \sum_{k=1}^{n} C_k\, \vec{w}_k.$$

According to Eq. (2.95), the solution is obviously asymptotically stable if and only if

$$\boxed{\mathrm{Re}\{\lambda_k\} < 0}\qquad\qquad(2.96)$$

holds for all $k \in \{1, 2, \ldots, n\}$, since only then does

$$\lim_{t\to\infty} \vec{r}(t) = 0$$

hold for arbitrary initial conditions. In this case, $\vec{r} = 0$ is an asymptotically stable fixed point.

Now we raise the question whether further fixed points exist. This is the case for

$$\vec{v}(\vec{r}) = A \cdot \vec{r} = 0$$

with

$$\vec{r} \neq 0,$$

i.e., only for

$$\det A = 0.$$

In Sect. 2.8.8, we will see that this is the condition for a degenerate, i.e., nonisolated, fixed point (see Definition 2.6, p. 59).

Let us now determine a map that transforms an initial value $\vec{r}_0$ into a vector $\vec{r}(t)$ that satisfies the linear autonomous system of ODEs

$$\frac{d\vec{r}}{dt} = \vec{v}(\vec{r}) = A \cdot \vec{r}.$$

The overall solution

$$\vec{r}(t) = \sum_{k=1}^{n} C_k\, \vec{w}_k e^{\lambda_k t}$$

with eigenvectors

$$\vec{w}_k = \begin{pmatrix} w_{k1} \\ w_{k2} \\ \cdots \\ w_{kn} \end{pmatrix}$$

may, due to

$$x_i(t) = \sum_{k=1}^{n} C_k \, w_{ki} \, e^{\lambda_k t},$$

be written as the matrix equation

$$\vec{r}(t) = \begin{pmatrix} w_{11} & w_{21} & \cdots & w_{n1} \\ w_{12} & w_{22} & \cdots & w_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ w_{1n} & w_{2n} & \cdots & w_{nn} \end{pmatrix} \cdot \begin{pmatrix} C_1 e^{\lambda_1 t} \\ C_2 e^{\lambda_2 t} \\ \cdots \\ C_n e^{\lambda_n t} \end{pmatrix}.$$

We define

$$X_A(0) = \begin{pmatrix} w_{11} & w_{21} & \cdots & w_{n1} \\ w_{12} & w_{22} & \cdots & w_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ w_{1n} & w_{2n} & \cdots & w_{nn} \end{pmatrix}$$

(matrix of the eigenvectors),

$$\vec{c} = \begin{pmatrix} C_1 \\ C_2 \\ \cdots \\ C_n \end{pmatrix},$$

and

$$D_A^{\lambda}(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{pmatrix}.$$

Hence we have

$$\vec{r}(t) = X_A(0) \cdot D_A^\lambda(t) \cdot \vec{c}.$$

For $t = 0$, we obtain

$$\vec{r}(0) = X_A(0) \cdot D_A^\lambda(0) \cdot \vec{c} = X_A(0) \cdot I \cdot \vec{c} = X_A(0) \cdot \vec{c}.$$

Due to

$$D_A^\lambda(0) = I,$$

the definition

$$X_A(t) = X_A(0) \cdot D_A^\lambda(t) \tag{2.97}$$

makes sense. Finally, we obtain

$$\vec{r}(t) = X_A(t) X_A(0)^{-1} \vec{r}(0).$$

Using the **matrix exponential function**

$$e^{tA} = X_A(t) X_A(0)^{-1}, \tag{2.98}$$

one also writes

$$\vec{r}(t) = e^{tA} \, \vec{r}(0). \tag{2.99}$$

This equation obviously determines the global flow (Guckenheimer [33, Eqn. (1.1.9), p. 9]) if the following definition is used:

Definition 2.11. A (global) **flow** is a continuous map $\Phi : \mathbb{R} \times D \rightarrow D$, which transforms each initial value $\vec{r}(0) = \vec{r}_0 \in D$ ($D \subset \mathbb{R}^n$ open) into a vector $\vec{r}(t)$ ($t \in \mathbb{R}$) satisfying the following conditions:

$$\Phi_0 = \text{id}, \quad \text{i.e.,} \quad \Phi_0(\vec{r}_0) = \vec{r}_0 \quad \text{for all } \vec{r}_0 \in D,$$

$$\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}, \quad \text{i.e.} \quad \Phi_{t_1+t_2}(\vec{r}_0) = \Phi_{t_1}(\Phi_{t_2}(\vec{r}_0)) \quad \text{for all } \vec{r}_0 \in D, \quad t_1, t_2 \in \mathbb{R}.$$

Here we have defined $\Phi_t(\vec{r}) := \Phi(t, \vec{r})$.

(Cf. Wiggins [38, Proposition 7.4.3, p. 93], Wirsching [31, Definition 8.6], Amann [39, p. 123/124].)

The interpretation of this definition is simple: If no time passes, one remains at the same point. Instead of moving from a first point to a second one in the time span

t_2 and then from this second one to a third one in a time span t_1, one may go directly from the first to the third in the time span $t_1 + t_2$.

Remark.

- A flow (also called a **phase flow**) is called a **global flow** if it is defined for all $t \in \mathbb{R}$ (as in Definition 2.11), a **semiflow** if it is defined for all $t \in \mathbb{R}^+$, and a **local flow** if it is defined for $t \in I$ (open interval I with $0 \in I$).
- For semiflows and local flows, Definition 2.11 has to be modified.
- In the modern mathematical literature, a dynamical system is *defined* as a flow. In our introduction, however, the dynamical system was initially described by an ODE, and the corresponding velocity vector field induced the flow.

Final remark: If the matrix A does not possess n linearly independent eigenvectors, then no diagonalization is possible, in which this case **generalized eigenvectors** may be used (cf. Guckenheimer [33, p. 9]). These are defined by

$$(A - \lambda I)^p \cdot \vec{w} = 0,$$
$$(A - \lambda I)^{p-1} \cdot \vec{w} \neq 0,$$

and may be used to transform any quadratic matrix A into **Jordan canonical form** (cf. Burg et al. [40, vol. II, p. 293] or Perko [30, Sect. 1.8]). As the formula shows, eigenvectors are also generalized eigenvectors (for $p = 1$).

2.8.7 Topological Orbit Equivalence

In this section, we shall define what it means to say that two vector fields are topologically orbit equivalent. Of course, the term **topological orbit equivalence** will include the case that the two vector fields can be transformed into each other by a simple rotation.

In order to simplify the situation even further, we assume that the two vector fields are given by
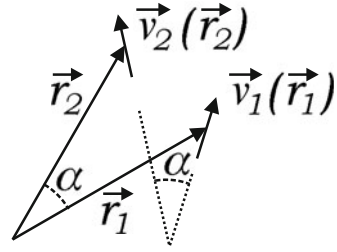
$$\vec{v}_1(\vec{r}_1) = A \cdot \vec{r}_1 \tag{2.100}$$

and

$$\vec{v}_2(\vec{r}_2) = B \cdot \vec{r}_2, \tag{2.101}$$

where A and B are quadratic matrices.

Fig. 2.5 Rotation of a vector
field

If one vector field can be obtained as a result of rotating the other one, there must
be a rotation matrix M such that

$$\vec{v}_2 = M \cdot \vec{v}_1 = M \cdot A \cdot \vec{r}_1 \tag{2.102}$$

holds. Now $\vec{v}_2$ depends on $\vec{r}_1$. In order to make $\vec{v}_2$ dependent on $\vec{r}_2$, we must rotate
the coordinates in the same way as the vector field (see Fig. 2.5):

$$\vec{r}_2 = M \cdot \vec{r}_1. \tag{2.103}$$

Hence, we obtain

$$\vec{v}_2 = M \cdot A \cdot M^{-1} \cdot \vec{r}_2.$$

Since M is invertible as a rotation matrix, the matrix

$$B = M \cdot A \cdot M^{-1}$$

describes the well-known similarity transformation that may also be written in the
form

$$B \cdot M = M \cdot A.$$

A **similarity transformation**, however, is usually written in the form

$$B = \tilde{M}^{-1} \cdot A \cdot \tilde{M},$$

so that we have to define $\tilde{M} = M^{-1}$ here.

Now we observe that two orbits can be identical even though the corresponding
solution curves are parameterized in a different way.

A flow for Eq. (2.100) will be denoted by

$$\vec{r}_1(t_1, \vec{r}_{10}),$$

and a flow for Eq. (2.101) by

$$\vec{r}_2(t_2, \vec{r}_{20}).$$

In order to transform the orbits of these flows into each other, the starting points must be mapped first:

$$\vec{r}_{20} = M \cdot \vec{r}_{10}.$$

Our requirement that different parameterizations be allowed for both solution curves may be translated as follows: For every time t_1, there is a time t_2 such that

$$\vec{r}_2(t_2, \vec{r}_{20}) = M \cdot \vec{r}_1(t_1, \vec{r}_{10}),$$

and therefore

$$\vec{r}_2(t_2, M \cdot \vec{r}_{10}) = M \cdot \vec{r}_1(t_1, \vec{r}_{10})$$

holds. This formula must be included in the general definition of topological orbit equivalence if rotations are to be allowed as topologically equivalent transformations.

The previous considerations make the following definition transparent:

Definition 2.12. Two C^1 vector fields $\vec{v}_1(\vec{r}_1)$ and $\vec{v}_2(\vec{r}_2)$ are called topologically orbit equivalent[15] if a homeomorphism h exists such that for every pair $\vec{r}_{10}, t_1$, there exists t_2 such that

$$\Phi_{t_2}^{\vec{v}_2}(h(\vec{r}_{10})) = h(\Phi_{t_1}^{\vec{v}_1}(\vec{r}_{10})) \tag{2.104}$$

holds. Here, the orientation of the orbits must be preserved. If in addition, the parameterization by time is preserved, the vector fields are called **topologically conjugate**. In this definition, $\Phi_t^{\vec{v}}$ denotes the flow that is induced by the vector field $\vec{v}$.

(Cf. Sastry [36, Definition 7.18, p. 303], Wiggins [38, Definition 19.12.1, p. 346], Guckenheimer [33, p. 38, Definition 1.7.3].)

---

[15]Sometimes, the abbreviation "TOE" is used for the property "topological orbit equivalence" (cf. Jackson [41]).

Fig. 2.6 Examples of
topological orbit equivalence
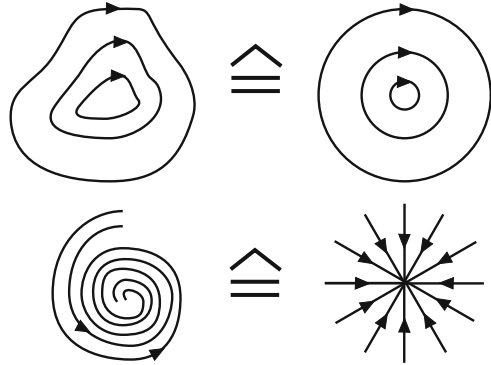
Remark. A **homeomorphism** is a continuous map whose inverse map exists and is also continuous. The fact that a homeomorphism is used as a generalization of the rotation matrix, used as an example above, leads to the following features:

- Not only linear maps are allowed, but also nonlinear ones.
- The requirement that the map be continuous guarantees that neighborhoods of a point are mapped to neighborhoods of its image point. Therefore, the orbits are deformed but not torn apart. Two examples for topological orbit equivalence are shown in Fig. 2.6.

The fact that the validity of Eq. (2.104) is required for each initial point ensures that all orbits are transformed into each other. Hence, the entire phase portraits will be equivalent.

The preservation of the orientation may be checked by means of a continuously differentiable function $t_2(\vec{r}_{10}, t_1)$ with $\frac{\partial t_2}{\partial t_1} > 0$ (cf. Perko [30, Sect. 3.1, Remark 2, p. 183/184]).

Please note that different authors use slightly different definitions. In cases of doubt, one should therefore check the relevant definitions thoroughly.

Let us now consider the case that a vector field $\vec{v}(\vec{r}) = A \cdot \vec{r}$ is given by a real $n \times n$ matrix A and that we want to check whether this vector field is topologically orbit equivalent to a simpler vector field. Often, diagonalization is possible. This case will be discussed in the following.

Remark.

- In case diagonalization is not possible, it is *always* possible to transform the matrix into Jordan canonical form.
- Diagonalization of an $n \times n$ matrix is possible if and only if for each eigenvalue, the **algebraic multiplicity** (multiplicity of the zeros of the characteristic polynomial) equals the **geometric multiplicity** (number of linearly independent eigenvectors).

- Diagonalization of an $n \times n$ matrix is possible if and only if it possesses n linearly independent eigenvectors.
- Diagonalization is possible for every symmetric matrix with real elements.

Consider a matrix A for which diagonalization is possible. We will show now that the diagonal matrix[16]

$$B = X_A(0)^{-1} \cdot A \cdot X_A(0) \tag{2.105}$$

does in fact lead to a topologically orbit equivalent vector field. Here, $\tilde{M} = M^{-1} = X_A(0)$ denotes the matrix of the n eigenvectors of A. These are linearly independent, since diagonalization of A is possible (cf. Burg/Haf/Wille [40, vol. II, p. 280, Theorem 3.52]).

According to Eqs. (2.98) and (2.99), the flows are given by

$$\Phi_{t_1}^{\vec{v}_1}(\vec{r}_{10}) = X_A(t_1) X_A(0)^{-1} \vec{r}_{10}$$

for A and by

$$\Phi_{t_2}^{\vec{v}_2}(\vec{r}_{20}) = X_B(t_2) X_B(0)^{-1} \vec{r}_{20}$$

for B. In our case, the homeomorphism h is given by the matrix M. We therefore obtain

$$\Phi_{t_2}^{\vec{v}_2}(h(\vec{r}_{10})) = \Phi_{t_2}^{\vec{v}_2}(M \cdot \vec{r}_{10}) = X_B(t_2) X_B(0)^{-1} X_A(0)^{-1} \cdot \vec{r}_{10}.$$

On the other hand,

$$h(\Phi_{t_1}^{\vec{v}_1}(\vec{r}_{10})) = M \cdot \Phi_{t_1}^{\vec{v}_1}(\vec{r}_{10}) = X_A(0)^{-1} X_A(t_1) X_A(0)^{-1} \cdot \vec{r}_{10}$$

holds. We see that these expressions are equal for every initial vector $\vec{r}_{10}$ if

$$X_B(t_2) X_B(0)^{-1} X_A(0)^{-1} = X_A(0)^{-1} X_A(t_1) X_A(0)^{-1},$$

or

$$X_B(t_2) X_B(0)^{-1} = X_A(0)^{-1} X_A(t_1),$$

---

[16]Please note that $X_A(0)$ is the matrix of eigenvectors of A such that Eq. (2.105) actually represents the diagonalization step.

is valid. Due to Eq. (2.97), we know that

$$X_A(t_1) = X_A(0) \cdot D_A^\lambda(t_1)$$

and

$$X_B(t_2) = X_B(0) \cdot D_B^\lambda(t_2)$$

hold, so that the equation

$$X_B(0) \cdot D_B^\lambda(t_2) X_B(0)^{-1} = X_A(0)^{-1} X_A(0) \cdot D_A^\lambda(t_1) = D_A^\lambda(t_1)$$

has to be verified. Since B is a diagonal matrix, the Cartesian unit vectors are eigenvectors, so that

$$X_B(0) = I$$

is valid. Therefore, we have only to check whether

$$D_B^\lambda(t_2) = D_A^\lambda(t_1)$$

is true. Since the eigenvalues of A and B are equal due to diagonalization, we obtain

$$D_A^\lambda(t_1) = D_B^\lambda(t_1).$$

Here we had only to set $t_2 = t_1$. We have shown that diagonalization leads to topologically orbit equivalent vector fields.

2.8.8 Classification of Fixed Points of an Autonomous Linear System of Second Order

The considerations presented above indicate that a similarity transformation

$$B = \tilde{M}^{-1} \cdot A \cdot \tilde{M}$$

always leads to topologically orbit equivalent vector fields. Every similarity transformation leaves the eigenvalues unchanged (cf. Burg/Haf/Wille [40, vol. II, p. 272, 3.17]). This leads us to the assumption that the eigenvalues of a matrix at least influence the topological properties of the related vector field. Therefore, the eigenvalues are now used to characterize the fixed points.

We calculate the eigenvalues of a two-dimensional matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

with constant real elements. This leads to

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

$$\Rightarrow \lambda^2 - \lambda(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21} = 0$$

$$\Rightarrow \lambda^2 - \lambda \, \text{tr}A + \det A = 0. \tag{2.106}$$

Hence, we obtain

$$\lambda = \frac{\text{tr}A}{2} \pm \sqrt{\frac{(\text{tr}A)^2}{4} - \det A} = B \pm \sqrt{C} \tag{2.107}$$

with

$$B = \frac{\text{tr}A}{2}, \qquad C = \frac{(\text{tr}A)^2}{4} - \det A. \tag{2.108}$$

We now try to distinguish as many cases as possible:

1. Both eigenvalues are real ($C \geq 0$).

   (a) Both are positive

      (i) $\lambda_1 > \lambda_2 > 0$
      (ii) $\lambda_1 = \lambda_2 > 0$

   (b) Both are negative

      (i) $\lambda_1 < \lambda_2 < 0$
      (ii) $\lambda_1 = \lambda_2 < 0$

   (c) One is positive, one is negative: $\lambda_1\lambda_2 < 0$
   (d) One equals 0 ($\lambda_1 = 0$):

      (i) $\lambda_2 > 0$
      (ii) $\lambda_2 < 0$
      (iii) $\lambda_2 = 0$

2. Imaginary eigenvalues: $C < 0$, $B = 0$, $\lambda_2 = -\lambda_1$
3. Complex eigenvalues: $C < 0$, $\lambda_1 = \lambda_2^*$

   (a) $B = \text{Re}\{\lambda\} < 0$
   (b) $B = \text{Re}\{\lambda\} > 0$

Hence we have found 11 distinct cases. If one eigenvalue is zero, then Eq. (2.106) leads to

$$\det A = 0.$$

As we will show now, this means in general that one row is a multiple of the other row, which contains the special case that one or both rows are zero.

If we now assume that the second row is a multiple of the first one,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ ka_{11} & ka_{12} \end{pmatrix},$$

we obtain the following condition for the eigenvalues:

$$(a_{11} - \lambda)(ka_{12} - \lambda) - ka_{11}a_{12} = 0$$
$$\Rightarrow \lambda^2 - \lambda(a_{11} + ka_{12}) = 0.$$

This shows that at least one eigenvalue is zero. The same can be shown for the case that the first equation is a multiple of the second one. Since one eigenvalue is 0, the relation

$$\det(A - \lambda I) = 0$$

leads to

$$\det A = 0.$$

In conclusion, the following statements for our two-dimensional case are equivalent:

- $\det A = 0$.
- One row of A is a multiple of the other row.
- At least one eigenvalue is 0.

The following general theorem holds:

Theorem 2.13. *The following statements for a quadratic matrix A are equivalent:*

- *The quadratic matrix A is regular.*
- *All row vectors (or column vectors) of A are linearly independent.*
- *$\det A \neq 0$.*
- *All eigenvalues of A are nonzero.*
- *A is invertible.*

In our two-dimensional case, the statement that for det $A = 0$, one row is a multiple of the other one means that the equation $\vec{v}(\vec{r}) = A \cdot \vec{r} = 0$ is satisfied along a line through the origin or even everywhere. Hence, we have an infinite number of fixed points that are not separated from each other. In this case, we speak of **degenerate fixed points** (see Definition 2.6 on p. 59). This case will henceforth be excluded[17] (case 1.d), so that the number of relevant cases is reduced from 11 to 8. According to the behavior of the vector field in the vicinity of the fixed point $\vec{r}_F = 0$, the fixed points are named as follows:

1. Both eigenvalues are real ($C \geq 0$).

   (a) Both positive

      (i) $\lambda_1 > \lambda_2 > 0$: **unstable node**
      (ii) $\lambda_1 = \lambda_2 > 0$: **unstable improper node or unstable star**

   (b) Both negative

      (i) $\lambda_1 < \lambda_2 < 0$: **stable node**
      (ii) $\lambda_1 = \lambda_2 < 0$: **stable improper node or stable star**

   (c) One positive, one negative: $\lambda_1 \lambda_2 < 0$: **saddle point**

2. Imaginary eigenvalues: $C < 0$, $B = 0$, $\lambda_2 = -\lambda_1$: **center** or **elliptic fixed point**
3. Complex eigenvalues: $C < 0$, $\lambda_1 = \lambda_2^*$

   (a) $\mathrm{Re}\{\lambda\} < 0$: **stable spiral point** or **stable focus**
   (b) $\mathrm{Re}\{\lambda\} > 0$: **unstable spiral point** or **unstable focus**

The classification of fixed points is summarized in Table 2.4 on p. 74 (cf. [29, 42, 43]). As the table shows, not only the eigenvalues can be used to characterize the fixed points. In the column "Topology", the numbers in parentheses denote the quantity of eigenvalues with positive real part. The column "Topology" also contains the so-called **index** in brackets. In order to calculate the index, one considers a closed path around the fixed point with an orientation that is mathematically positive. Now one checks how many revolutions the vectors of the vector field perform while "walking" on the path. If, e.g., the vectors of the vector field also perform one revolution in mathematically positive orientation, the index is $+1$. If the vector field rotates in opposite direction, the index is -1.

Figures 2.7, 2.8, 2.9, 2.10, 2.11, and 2.12 on p. 75 and 76 show how orbits in the vicinity of the fixed point look in principle for each type of fixed point. In case the fixed point is a stable node, star, or spiral point, the orientation of the solution curves will be towards the fixed point in the middle; in case of an unstable node, star, or spiral point, all solution curves will be directed outwards. Each picture is just an example; in the specific case under consideration the orbits may of course be deformed significantly.

---

[17]Please note that small changes of the entries of the matrix A are likely to convert degenerate fixed points into nondegenerate (isolated) fixed points (cf. Sastry [36, p. 2]).

**Table 2.4** Classification of isolated fixed points in the plane (regular Jacobian matrix)

	Stability	**Topology**	**Eigenvalues**	**Fixed Point**
$\text{Re}\{\lambda\} = 0$	stable	center [+1] (-)	imaginary $\lambda_1 = -\lambda_2$	center =elliptic fixed point
hyperbolic $\text{Re}\{\lambda\} \neq 0$	asymptotically stable	sink [+1] (0) $\text{Re}\{\lambda\} < 0$	real $\lambda_1 < \lambda_2 < 0$	stable improper node (two[a] tangents)
			real $\lambda_1 = \lambda_2 < 0$	stable improper node (one tangent, one eigenvector)
				stable proper node =stable star (many tangents, two eigenvectors)
			complex $\text{Re}\{\lambda\} < 0$	stable spiral point =stable focus
	unstable	source [+1] (2) $\text{Re}\{\lambda\} > 0$	real $\lambda_1 > \lambda_2 > 0$	unstable improper node (two[a] tangents)
			real $\lambda_1 = \lambda_2 > 0$	unstable improper node (one tangent, one eigenvector)
				unstable proper node =unstable star (many tangents, two eigenvectors)
			complex $\text{Re}\{\lambda\} > 0$	unstable spiral point =unstable focus
		saddle point [-1] (1)	real $\lambda_1 < 0 < \lambda_2$	saddle point

[a]one tangent for two orbits, one tangent for all other orbits
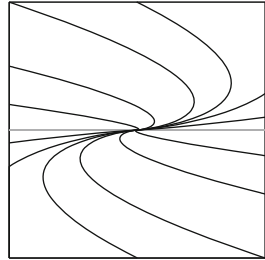
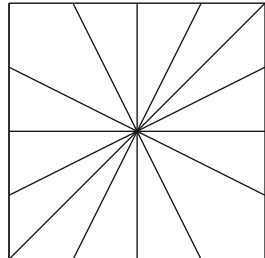**Fig. 2.7** Center
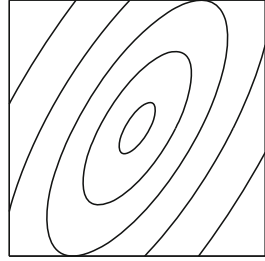
Fig. 2.8 Node with two tangents
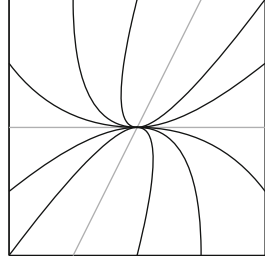
Fig. 2.9 Node with one tangent
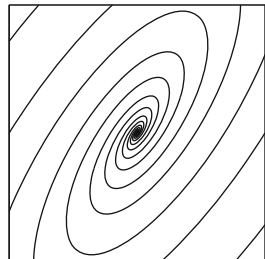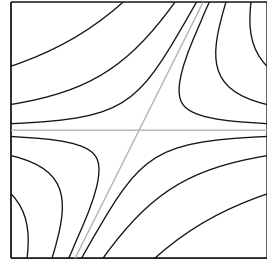

Fig. 2.10 Star

Fig. 2.11 Spiral point

Fig. 2.12 Saddle point

2.8.9   Nonlinear Systems

Consider the nonlinear autonomous system

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}) \tag{2.109}$$

with the initial condition

$$\vec{r}(0) = \vec{r}_0,$$

where $\vec{v}(\vec{r}) \in C^2$ has a fixed point $\vec{r}_\mathrm{F}$ with

$$\vec{v}(\vec{r}_\mathrm{F}) = 0.$$

As the results summarized in this section show, the **linearization**

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = A \cdot \vec{r} \tag{2.110}$$

of the system (2.109), where $A = D\vec{v}(\vec{r}_\mathrm{F})$ is the **Jacobian matrix**[18] at $\vec{r} = \vec{r}_\mathrm{F}$, is a powerful tool for analyzing a nonlinear system in the vicinity of its fixed points.

---

[18] The Jacobian matrix is defined as

$$D\vec{v}(\vec{r}) := \frac{\partial \vec{v}}{\partial \vec{r}} := \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \cdots & \frac{\partial v_1}{\partial x_n} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \cdots & \frac{\partial v_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial x_1} & \frac{\partial v_n}{\partial x_2} & \cdots & \frac{\partial v_n}{\partial x_n} \end{pmatrix}$$

where $\vec{r} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots + x_n \vec{e}_n$.

The determinant of the Jacobian matrix is called **Jacobian determinant** or simply **Jacobian**, and one writes

$$\det D\vec{v}(\vec{r}) = \frac{\partial(v_1, v_2, \ldots, v_n)}{\partial(x_1, x_2, \ldots, x_n)}.$$

Please note that $\vec{r} = 0$ in Eq. (2.110) corresponds to $\vec{r} = \vec{r}_F$ in Eq. (2.109), i.e., the fixed point of the nonlinear system was shifted to the origin of the linearized system.

Theorem 2.14. *Consider the nonlinear system (2.109) with the linearization (2.110) at a fixed point $\vec{r}_F$. If A is nonsingular, then the fixed point $\vec{r}_F$ is isolated (i.e., nondegenerate).*

(See Sastry [36, Proposition 1.5, p. 13], Perko [30, Definition 2, p. 173].)

If one or more eigenvalues of the Jacobian matrix are zero, the fixed point is a degenerate fixed point. This is the generalization of the linear case.

Definition 2.15. A fixed point is called a hyperbolic fixed point if no eigenvalue of the Jacobian matrix has zero real part.

Theorem 2.16. *If the fixed point $\vec{r}_F$ is a hyperbolic fixed point, then there exist two neighborhoods U of $\vec{r}_F$ and V of $\vec{r} = 0$ and a homeomorphism $h : U \rightarrow V$, such that h transforms the orbits of Eq. (2.109) into orbits of*

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = A \cdot \vec{r} \qquad with \qquad A = D\vec{v}(\vec{r}_F).$$

Orientation and parameterization by time are preserved.

(Cf. Guckenheimer [33, Theorem 1.3.1 (Hartman-Grobman), p. 13], Perko [30, Theorem Sect. 2.8, p. 120], and Bronstein [44, Sect. 11.3.2].)

In other words, we may state the following theorem.

Theorem 2.17 (Hartman–Grobman). *Let $\vec{r}_F$ be a hyperbolic fixed point. Then the nonlinear problem*

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}), \qquad \vec{v} \in C^1(D), \qquad D \subset \mathbb{R}^n \ open,$$

and the linearized problem

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = A \cdot (\vec{r} - \vec{r}_F)$$

with

$$A = D\vec{v}(\vec{r}_F)$$

are topologically conjugate in a neighborhood of $\vec{r}_F$.

(See Wiggins [38, Theorem 19.12.6, p. 350].)

If the fixed point is not hyperbolic, i.e., a center (elliptic fixed point), then the smallest nonlinearities are sufficient to create a stable or an unstable spiral point. This is why the theorem refers to hyperbolic fixed points only.

If the real parts of all eigenvalues of $D\vec{v}(\vec{r}_F)$ are negative, then $\vec{r}_F$ is asymptotically stable. If the real part of at least one eigenvalue is positive, then $\vec{r}_F$ is unstable:

Theorem 2.18. *Let $D \subset \mathbb{R}^n$ be an open set, $\vec{v}(\vec{r})$ continuously differentiable on D, and $\vec{r}_F$ a fixed point of*

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \vec{v}(\vec{r}).$$

If the real parts of all eigenvalues of $D\vec{v}(\vec{r}_F)$ are negative, then $\vec{r}_F$ is asymptotically stable. If $\vec{r}_F$ is stable, then no eigenvalue has positive real part.

(Cf. Bronstein [44, Sect. 11.3.1], Perko [30, Theorem 2, p. 130].)

A saddle point has the special property that two trajectories exist that approach the saddle point for $t \to \infty$, whereas two different trajectories exist that approach the saddle point for $t \to -\infty$ (cf. [30, Sect. 2.10, Definition 5]). These four trajectories define a **separatrix**. Loosely speaking, a separatrix is a trajectory that "meets" the saddle point.

2.8.10   Characteristic Equation

Consider the autonomous linear homogeneous nth-order ordinary differential equation

$$a_n \frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) = 0. \qquad (2.111)$$

As usual (see Sect. 2.8.1), we define the vector $\vec{r} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ by

$$x_1 = y,$$
$$x_2 = \frac{\mathrm{d}y}{\mathrm{d}t},$$
$$\cdots$$
$$x_n = \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}},$$

which leads to

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = x_3,$$
$$\cdots$$
$$\dot{x}_n = -\frac{a_0}{a_n}x_1 - \frac{a_1}{a_n}x_2 - \cdots - \frac{a_{n-1}}{a_n}x_n,$$

in order to obtain the standard form

$$\frac{d\vec{r}}{dt} = \vec{v}(\vec{r}).$$

This may be written as

$$\frac{d\vec{r}}{dt} = A \cdot \vec{r}$$

if the following $n \times n$ matrix is defined:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & -\frac{a_3}{a_n} & \cdots & -\frac{a_{n-2}}{a_n} & -\frac{a_{n-1}}{a_n} \end{pmatrix}. \tag{2.112}$$

According to Sect. 2.8.6, Eq. (2.96), we know that asymptotic stability is reached if all eigenvalues of this system matrix have negative real part. Therefore, we now describe how to find the eigenvalues based on the requirement that the determinant

$$D_n^F = \det(A - \lambda I) \tag{2.113}$$

equal zero. Let us begin with $n = 2$ as an example:

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{a_0}{a_2} & -\frac{a_1}{a_2} \end{pmatrix}$$

$$\Rightarrow D_2^F = \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -\frac{a_0}{a_2} & -\frac{a_1}{a_2} - \lambda \end{vmatrix} = \lambda^2 + \frac{a_1}{a_2}\lambda + \frac{a_0}{a_2} \stackrel{!}{=} 0.$$

For $n = 3$, we obtain

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{a_0}{a_3} & -\frac{a_1}{a_3} & -\frac{a_2}{a_3} \end{pmatrix}$$

$$\Rightarrow D_3^F = \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -\frac{a_0}{a_3} & -\frac{a_1}{a_3} & -\frac{a_2}{a_3} - \lambda \end{vmatrix} = -\lambda^3 - \frac{a_2}{a_3}\lambda^2 - \frac{a_1}{a_3}\lambda - \frac{a_0}{a_3} \stackrel{!}{=} 0.$$

These two results lead us to the assumption that

$$D_n^F = (-1)^n \left(\lambda^n + \frac{a_{n-1}}{a_n}\lambda^{n-1} + \frac{a_{n-2}}{a_n}\lambda^{n-2} + \cdots + \frac{a_2}{a_n}\lambda^2 + \frac{a_1}{a_n}\lambda + \frac{a_0}{a_n} \right)$$

$$= (-1)^n \sum_{k=0}^{n} \frac{a_k}{a_n}\lambda^k \tag{2.114}$$

holds in general. In Appendix A.6, it is shown that this is indeed true. The requirement that the polynomial in Eq. (2.114) equal zero is called the **characteristic equation** of the ODE (2.111). One easily sees that the characteristic equation is also obtained if the Laplace transform is applied to the original ODE (2.111):

$$\left(a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \right) Y(s) = 0$$

$$\Rightarrow a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 = 0 \qquad \text{for} \quad Y(s) \neq 0.$$

The matrix A is called the **Frobenius companion matrix** of the polynomial. Please note that instead of finding the zeros (roots) of the polynomial, one may also determine the eigenvalues of the companion matrix A and vice versa.

Hence, asymptotic stability of the dynamical system defined by the ODE (2.111) is equivalently shown

- if all zeros of the characteristic equation have negative real part.
- if all eigenvalues of the system matrix have negative real part.

In case of asymptotic stability, one also calls the system matrix a **strictly** or **negative stable matrix** (cf. [45, Definition 2.4.2]) or a **Hurwitz matrix**.

2.9 Continuity Equation

Consider a particle density ρ in space and a velocity field $\vec{v}$ that moves the particles. We will now calculate how the particle density in a fixed volume V changes due to the velocity field.

For this purpose, we consider a small volume element ΔV at the surface of the three-dimensional domain V. As shown in Fig. 2.13, this contains in total
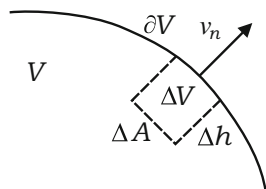
$$\Delta n = \Delta V \, \rho = \Delta h \, \Delta A \, \rho$$

particles. During the time interval Δt, this quantity of

$$\Delta n = v_n \Delta t \, \Delta A \, \rho = \rho \, \Delta t \, \vec{v} \cdot \Delta \vec{A}$$

Fig. 2.13 Volume element at
the surface of a region

particles will leave the domain V, where $v_n = \Delta h / \Delta t$ denotes the normal component of the velocity vector $\vec{v}$ with respect to the surface of the domain V. Hence, we have

$$\int_V \rho(t + \Delta t) \, \mathrm{d}V = \int_V \rho(t) \, \mathrm{d}V - \oint_{\partial V} \rho \Delta t \, \vec{v} \cdot \mathrm{d}\vec{A}.$$

As a limit for $\Delta t \to 0$, one therefore obtains

$$\int_V \dot{\rho} \, \mathrm{d}V = -\oint_{\partial V} \rho \, \vec{v} \cdot \mathrm{d}\vec{A}.$$

According to Gauss's theorem,

$$\oint_{\partial V} \vec{V} \cdot \mathrm{d}\vec{A} = \int_V \operatorname{div} \vec{V} \, \mathrm{d}V,$$

one concludes by setting $\vec{V} = \rho \, \vec{v}$:

$$\int_V \dot{\rho} \, \mathrm{d}V = -\int_V \operatorname{div}(\rho \, \vec{v}) \, \mathrm{d}V.$$

Since this equation must be valid for arbitrary choices of the domain V, one obtains

$$\boxed{-\dot{\rho} = \operatorname{div}(\rho \, \vec{v}).} \tag{2.115}$$

This is the **continuity equation**, for which we only assumed that no particles disappear and no particles are generated. Instead of the particle density, one could have considered different densities, such as the mass density, assuming mass conservation in that case. If we take the charge density as an example, charge conservation leads to

$$\boxed{-\dot{\rho}_q = \operatorname{div}(\rho_q \, \vec{v}) = \operatorname{div} \vec{J},}$$

which we already know as Eq. (2.46) and where $\vec{J} = \rho_q \vec{v}$ is the **convection current density**.

Remark. If $\dot{\rho}_q = 0$ holds, then the density will remain constant at every location; one obtains a stationary flow with

$$\text{div}(\rho_q\,\vec{v}) = 0 \qquad \text{or} \quad \text{div}\,\vec{J} = 0.$$

This equation is known in electromagnetism for steady currents.

2.10 Area Preservation in Phase Space

In this section, we discuss how an area or a volume that is defined by the contained particles is modified when the particles are moving.

2.10.1 Velocity Vector Fields

Consider an arbitrary domain A in $\mathbb{R}^2$ at time t. Particles located inside the domain and on its boundary at time t will move a bit farther during the time span Δt. This movement is determined by the velocity field $\vec{v}(x, y)$.

Let us define a parameterization of the domain such that $x(\alpha, \beta)$ and $y(\alpha, \beta)$ are given depending on the parameters α and β. This leads to the area

$$A(t) = \int_A \mathrm{d}A = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} \left| \frac{\partial(x, y)}{\partial(\alpha, \beta)} \right| \, \mathrm{d}\alpha \, \mathrm{d}\beta.$$

The coordinates $\vec{r} = (x, y)$ denote each point of A. Such a point $\vec{r}$ will move to the new point

$$\vec{r}\,' = \vec{r} + \vec{v}(\vec{r})\Delta t$$

after the time span Δt. Since $\vec{r}$ depends on α and β, it follows that $\vec{r}\,'$ will also depend on these parameters. For the area of the deformed domain at the time $t + \Delta t$, we therefore get

$$A(t + \Delta t) = A' = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} \left| \frac{\partial(x', y')}{\partial(\alpha, \beta)} \right| \, \mathrm{d}\alpha \, \mathrm{d}\beta$$

with

$$x' = x + v_x\Delta t,$$
$$y' = y + v_y\Delta t.$$

Using the abbreviation

$$\xi = \frac{\partial(x, y)}{\partial(\alpha, \beta)} = \begin{vmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial \beta} \\ \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial \beta} \end{vmatrix} = \det(\vec{r}_\alpha, \vec{r}_\beta)$$

or

$$\xi' = \det(\vec{r}\,'_\alpha, \vec{r}\,'_\beta)$$

leads to

$$\xi' = \det(\vec{r}_\alpha + \vec{v}_\alpha \Delta t, \vec{r}_\beta + \vec{v}_\beta \Delta t) = \det(\vec{r}_\alpha, \vec{r}_\beta + \vec{v}_\beta \Delta t) + \det(\vec{v}_\alpha \Delta t, \vec{r}_\beta + \vec{v}_\beta \Delta t) =$$
$$= \det(\vec{r}_\alpha, \vec{r}_\beta) + \det(\vec{r}_\alpha, \vec{v}_\beta \Delta t) + \det(\vec{v}_\alpha \Delta t, \vec{r}_\beta) + \det(\vec{v}_\alpha \Delta t, \vec{v}_\beta \Delta t) =$$
$$= \xi + \Delta t \left(\det(\vec{r}_\alpha, \vec{v}_\beta) + \det(\vec{v}_\alpha, \vec{r}_\beta) \right) + \Delta t^2 \det(\vec{v}_\alpha, \vec{v}_\beta).$$

One obtains

$$\frac{\partial \xi}{\partial t} = \lim_{\Delta t \to 0} \frac{\xi' - \xi}{\Delta t} = \det(\vec{r}_\alpha, \vec{v}_\beta) + \det(\vec{v}_\alpha, \vec{r}_\beta) =$$

$$= \frac{\partial x}{\partial \alpha} \frac{\partial v_y}{\partial \beta} - \frac{\partial y}{\partial \alpha} \frac{\partial v_x}{\partial \beta} + \frac{\partial v_x}{\partial \alpha} \frac{\partial y}{\partial \beta} - \frac{\partial v_y}{\partial \alpha} \frac{\partial x}{\partial \beta} =$$

$$= \frac{\partial x}{\partial \alpha} \left(\frac{\partial v_y}{\partial x} \frac{\partial x}{\partial \beta} + \frac{\partial v_y}{\partial y} \frac{\partial y}{\partial \beta} \right) - \frac{\partial y}{\partial \alpha} \left(\frac{\partial v_x}{\partial x} \frac{\partial x}{\partial \beta} + \frac{\partial v_x}{\partial y} \frac{\partial y}{\partial \beta} \right) +$$

$$+ \frac{\partial y}{\partial \beta} \left(\frac{\partial v_x}{\partial x} \frac{\partial x}{\partial \alpha} + \frac{\partial v_x}{\partial y} \frac{\partial y}{\partial \alpha} \right) - \frac{\partial x}{\partial \beta} \left(\frac{\partial v_y}{\partial x} \frac{\partial x}{\partial \alpha} + \frac{\partial v_y}{\partial y} \frac{\partial y}{\partial \alpha} \right) =$$

$$= \frac{\partial v_y}{\partial y} \left(\frac{\partial x}{\partial \alpha} \frac{\partial y}{\partial \beta} - \frac{\partial x}{\partial \beta} \frac{\partial y}{\partial \alpha} \right) + \frac{\partial v_x}{\partial x} \left(\frac{\partial x}{\partial \alpha} \frac{\partial y}{\partial \beta} - \frac{\partial x}{\partial \beta} \frac{\partial y}{\partial \alpha} \right)$$

$$\Rightarrow \frac{\partial \xi}{\partial t} = \xi \operatorname{div} \vec{v}.$$

Since $\xi \neq 0$ is valid ($\xi = |\xi| \operatorname{sgn} \xi$, $\operatorname{sgn} \xi$ constant), one gets

$$\frac{\partial |\xi|}{\partial t} = |\xi| \operatorname{div} \vec{v}.$$

Due to

$$A(t) = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} |\xi| \, d\alpha \, d\beta,$$

one obtains

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} \frac{\partial |\xi|}{\partial t} \, \mathrm{d}\alpha \, \mathrm{d}\beta = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} |\xi| \, \operatorname{div} \vec{v} \, \mathrm{d}\alpha \, \mathrm{d}\beta.$$

Now it is obvious that the area remains constant for div $\vec{v} = 0$. If we were talking here about a fluid, such a fluid would obviously be **incompressible**; were one to try to compress it, the shape would be modified, but the total area (or volume) occupied by the particles would remain the same.

2.10.2 Maps

Now we analyze in a more general way how an area

$$A = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} \left| \frac{\partial (x, y)}{\partial (\alpha, \beta)} \right| \, \mathrm{d}\alpha \, \mathrm{d}\beta$$

is modified by a map

$$\vec{r}\,' = \vec{F}(\vec{r}),$$

which transforms each vector $\vec{r} = (x, y)$ into a vector $\vec{r}\,' = (x', y')$. The parameterization will remain the same. Each point of the domain moves to a new point, so that the shape of the domain will change in general. Hence, we have to calculate

$$A' = \int_{\beta_{\min}}^{\beta_{\max}} \int_{\alpha_{\min}}^{\alpha_{\max}} \left| \frac{\partial (x', y')}{\partial (\alpha, \beta)} \right| \, \mathrm{d}\alpha \, \mathrm{d}\beta.$$

According to Appendix A.5, we have

$$A' = A$$

if

$$\left| \frac{\partial (x', y')}{\partial (x, y)} \right| = 1$$

is satisfied; the Jacobian of area-preserving maps is obviously $+1$ or -1.

We now check this general formula for the situation discussed in the previous section, where a special map

$$\vec{r}\,' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \vec{r} + \vec{v} \, \Delta t = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix} \Delta t$$

was given. The Jacobian is then

$$
\begin{vmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 + \frac{\partial v_x}{\partial x}\,\Delta t & \frac{\partial v_x}{\partial y}\,\Delta t \\ \frac{\partial v_y}{\partial x}\,\Delta t & 1 + \frac{\partial v_y}{\partial y}\,\Delta t \end{vmatrix} = 1 + \Delta t \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \right)
$$

$$
+ \Delta t^2 \left(\frac{\partial v_x}{\partial x} \frac{\partial v_y}{\partial y} - \frac{\partial v_x}{\partial y} \frac{\partial v_y}{\partial x} \right).
$$

If one now wants to calculate

$$
\frac{\partial \xi}{\partial t} = \lim_{\Delta t \to 0} \frac{\xi' - \xi}{\Delta t},
$$

one obtains, due to

$$
\xi' = \frac{\partial (x', y')}{\partial (x, y)}\,\xi
$$

(see Appendix A.5), the relation

$$
\frac{\partial \xi}{\partial t} = \xi \lim_{\Delta t \to 0} \frac{\frac{\partial (x', y')}{\partial (x, y)} - 1}{\Delta t} = \xi \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \right) = \xi \,\mathrm{div}\,\vec{v},
$$

as above.

2.10.3   Liouville's Theorem

The statement derived above that the condition

$$
\boxed{\mathrm{div}\,\vec{v} = 0}
$$

leads to **area preservation** or—depending on the dimension—to **volume preservation** in phase space is called **Liouville's theorem**. This equation is also given as the condition for **incompressible flows**. Please note that one can speak of area preservation only if the area is defined in a unique way. This is, for example possible, if a continuous particle density ρ with clear boundaries in phase space is assumed, but not for a discrete distribution of individual particles (or only approximately if large numbers of particles are present). We will return to this problem later.

Liouville's theorem (and therefore also area/volume preservation) is also valid if

$$
\vec{v}(\vec{r}, t)
$$

depends explicitly on time (cf. Szebehely [46, p. 55], Fetter [47, p. 296], or Budó [48, p. 446]).

Fig. 2.14 Spring–mass
system
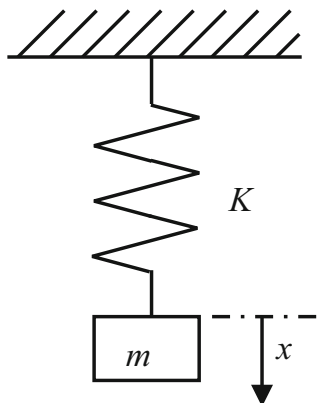
2.11 Hamiltonian Systems

Hamiltonian theory is usually developed in the scope of classical mechanics after
introduction of the Lagrangian formulation (cf. [19]). Here we choose a different
approach by introducing Hamiltonian functions directly. This can, of course, be no
replacement for intense studies of Hamiltonian mechanics, but it is sufficient to
understand some basics that are relevant in the following chapters of this book.

2.11.1 Example for Motivation

Consider the system sketched in Fig. 2.14. The spring constant K and the mass m
are known. The force balance leads to

$$m\ddot{x} = -Kx$$

$$\Leftrightarrow \ddot{x} + \frac{K}{m} x = 0. \qquad (2.116)$$

The general solution is obtained using the following ansatz:

$$x = A \, \cos(\omega t) + B \, \sin(\omega t),$$

$$\dot{x} = -A\omega \, \sin(\omega t) + B\omega \, \cos(\omega t),$$

$$\ddot{x} = -A\omega^2 \, \cos(\omega t) - B\omega^2 \, \sin(\omega t).$$

We obviously obtain

$$\omega^2 = \frac{K}{m}.$$

Fig. 2.15 Trajectory of the
spring–mass system
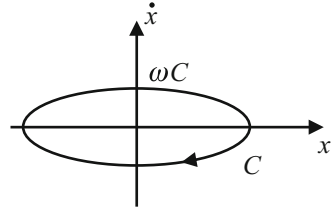
One can alternatively write x in the form

$$x = C \ \cos(\omega t - \varphi) = C \ \cos(\omega t) \ \cos \varphi + C \ \sin(\omega t) \ \sin \varphi. \qquad (2.117)$$

This leads to:

$$A = C \ \cos \varphi,$$
$$B = C \ \sin \varphi.$$

For $\dot{x}$ and $\ddot{x}$ one obtains

$$\dot{x} = -C\omega \ \sin(\omega t - \varphi), \qquad (2.118)$$

$$\ddot{x} = -C\omega^2 \ \cos(\omega t - \varphi) = -\omega^2 x. \qquad (2.119)$$

The result may be drawn as shown in Fig. 2.15.

The quantity φ obviously determines only the initial conditions, whereas C is the oscillation amplitude and thus characterizes the energy of the system.

The quantity C (regarded as a system property), as well as the energy W, remains constant on the trajectory. If, in general, we have an invariant H that depends on two variables q and p, then the trajectory $(q(t), p(t))$ will have the property that

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0$$

holds. One concludes that

$$\frac{\partial H}{\partial q} \frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial H}{\partial p} \frac{\mathrm{d}p}{\mathrm{d}t} = 0$$

$$\Leftrightarrow \frac{\partial H}{\partial q} \dot{q} + \frac{\partial H}{\partial p} \dot{p} = 0.$$

This equation is obviously satisfied if the following system of equations is valid:

$$\frac{\partial H}{\partial p} = \dot{q}, \qquad (2.120)$$

$$\frac{\partial H}{\partial q} = -\dot{p}. \tag{2.121}$$

These equations are called **Hamilton's equations**. The function $H(q, p)$ is called the **Hamiltonian**. We will now check whether this system of equations is actually satisfied in our example.

It is clear that the total energy of the system remains constant:

$$W(x, v) = \frac{K}{2}\, x^2 + \frac{1}{2}mv^2.$$

This can also be seen formally if the differential equation

$$m\,\ddot{x} + K\,x = 0$$

is multiplied by $\dot{x}$:

$$m\,\dot{x}\,\ddot{x} + K\,\dot{x}\,x = 0$$

$$\Leftrightarrow \frac{m}{2}\frac{d}{dt}(\dot{x}^2) + \frac{K}{2}\frac{d}{dt}(x^2) = 0$$

$$\Leftrightarrow \frac{dW}{dt} = 0.$$

Here we obviously have

$$\frac{\partial W}{\partial x} = Kx,$$

$$\frac{\partial W}{\partial v} = mv,$$

$$\dot{x} = v,$$

$$\dot{v} = \ddot{x} = -\frac{K}{m}x.$$

As a result, one obtains

$$\frac{\partial W}{\partial x} = -m\dot{v},$$

$$\frac{\partial W}{\partial v} = m\dot{x}.$$

These equations are not yet equivalent to the above-mentioned Hamilton equations. However, it is not a big step to work with $p = mv$ instead of v:

$$\frac{\partial W}{\partial x} = -\dot{p},$$

$$\frac{\partial W}{\partial p} = \frac{1}{m}\frac{\partial W}{\partial v} = \dot{x}.$$

Now the equations actually have the desired form; Hamilton's equations are satisfied. The function $W(x, p)$ is called a Hamiltonian, since it satisfies Hamilton's equations.

As shown above, C is also constant along the trajectory. We can obviously determine C as follows, based on Eqs. (2.117) and (2.118):

$$(\omega C)^2 = (\omega x)^2 + \dot{x}^2.$$

It seems to be useful to define the following quantities in order to get $C = \sqrt{\bar{q}^2 + \bar{p}^2}$:

$$\bar{q} = x, \qquad \bar{p} = \frac{\dot{x}}{\omega}.$$

Calculating the partial derivatives leads to

$$\frac{\partial C}{\partial \bar{p}} = \frac{1}{2C}\, 2\bar{p},$$

$$\frac{\partial C}{\partial \bar{q}} = \frac{1}{2C}\, 2\bar{q},$$

$$\dot{\bar{q}} = \dot{x} = \omega\bar{p}.$$

With the help of Eq. (2.116), one obtains

$$\dot{\bar{p}} = \frac{\ddot{x}}{\omega} = -\omega x = -\omega\bar{q}.$$

We therefore get two coupled differential equations:

$$\frac{\partial C}{\partial \bar{p}} = \frac{1}{\omega C}\, \dot{\bar{q}} \quad \Rightarrow \omega C\, \frac{\partial C}{\partial \bar{p}} = \dot{\bar{q}},$$

$$\frac{\partial C}{\partial \bar{q}} = -\frac{1}{\omega C}\, \dot{\bar{p}} \quad \Rightarrow \omega C\, \frac{\partial C}{\partial \bar{q}} = -\dot{\bar{p}}.$$

This is reminiscent of the product rule

$$\frac{\partial (C^2)}{\partial \bar{p}} = 2C\frac{\partial C}{\partial \bar{p}}, \qquad \text{or} \qquad \frac{\partial (C^2)}{\partial \bar{q}} = 2C\frac{\partial C}{\partial \bar{q}}.$$

If we therefore set

$$H = C^2 \frac{\omega}{2},$$

we again obtain Hamilton's equations:

$$\frac{\partial H}{\partial \bar{p}} = \dot{\bar{q}},$$

$$\frac{\partial H}{\partial \bar{q}} = -\dot{\bar{p}}.$$

We conclude that on the trajectory, the Hamiltonian

$$H(\bar{q}, \bar{p}) = \frac{\omega}{2} \bar{q}^2 + \frac{\omega}{2} \bar{p}^2$$

is constant if in our example, the generalized coordinate

$$\bar{q} = x$$

and the generalized momentum

$$\bar{p} = \dot{x}/\omega$$

are used. In our special case, $\bar{q}$ is a physical coordinate, but $\bar{p}$ is not the physical momentum. In general, $\bar{q}$ also does not need to be a physical coordinate. This explains the terms "generalized coordinate" and "generalized momentum." Here they formally play a similar mathematical role. We summarize:

- $H(q, p)$ is called a Hamiltonian if Hamilton's equations (2.120) and (2.121) are satisfied.
- The Hamiltonian describes a dynamical system.
- The quantities q and p are called a **generalized coordinate** and **generalized momentum**, respectively. They do not necessarily have to be identical to the physical coordinates and momenta.
- Different Hamiltonians may exist for the same dynamical system (in our example, $\frac{\omega}{2} C^2$ and W), and also different definitions of q and p are possible.

2.11.2   Arbitrary Number of Variables

Our introductory example contained only one coordinate and one momentum variable. For an arbitrary number of coordinate variables, Hamilton's equations are

$$\frac{\partial H}{\partial p_i} = \dot{q}_i, \qquad (2.122)$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i. \qquad (2.123)$$

In this case, the Hamiltonian

$$H(q_i, p_i, t)$$

depends on n generalized coordinates q_i ($1 \leq i \leq n$), on n generalized momentum variables p_i, and in general, explicitly on the time t. Its total derivative with respect to time is

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{i=1}^{n} \left[\frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i \right] + \frac{\partial H}{\partial t}.$$

By means of Eqs. (2.122) and (2.123), one obtains

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}.$$

This shows that if the Hamiltonian does not explicitly depend on time (as in our introductory example), it is constant along the trajectory. In contrast to this case, an explicit time dependence directly determines the time dependence along the trajectory.

Autonomous Hamiltonian systems $H(q_i, p_i)$ with no explicit time dependence are conservative systems, because H does not change with time (i.e., along the trajectory).

2.11.3   Flow in Phase Space

In general, we consider a system with n degrees of freedom. In this case, we have n generalized coordinates q_i and n generalized momentum variables p_i ($i \in \{1, 2, \ldots, n\}$).

The $2n$-dimensional space that is generated by these variables is called the phase space. If the $2n$ variables q_i and p_i are given at a time t_0, the system state is determined completely, and $q_i(t)$, $p_i(t)$ can be calculated for arbitrary times t (in the maximal interval of existence; see Sects. 2.8.3.3 and 2.8.3.4).

In order to show this, we combine coordinate and momentum variables as follows:

$$\vec{r} = \begin{pmatrix} q_1 \\ q_2 \\ \cdots \\ q_n \\ p_1 \\ p_2 \\ \cdots \\ p_n \end{pmatrix}, \qquad \vec{v} = \dot{\vec{r}} = \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \cdots \\ \dot{q}_n \\ \dot{p}_1 \\ \dot{p}_2 \\ \cdots \\ \dot{p}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p_1} \\ \frac{\partial H}{\partial p_2} \\ \cdots \\ \frac{\partial H}{\partial p_n} \\ -\frac{\partial H}{\partial q_1} \\ -\frac{\partial H}{\partial q_2} \\ \cdots \\ -\frac{\partial H}{\partial q_n} \end{pmatrix}.$$

In the last step, we used Hamilton's equations (2.122) and (2.123). Based on this definition, the problem has the standard form (2.94) of a dynamical system (see p. 50). We obtain

$$\text{div } \vec{v} = \sum_{k=1}^{n} \left(\frac{\partial^2 H}{\partial p_k \partial q_k} - \frac{\partial^2 H}{\partial q_k \partial p_k} \right) = 0.$$

Therefore, the flow in phase space corresponds to an incompressible fluid. Thus, Liouville's theorem is valid automatically, stating that the area/volume in phase space remains constant. We have assumed only the preservation of the number of particles and the validity of Hamilton's equations.

Liouville's theorem (and area/volume preservation) is also valid if the Hamiltonian

$$H(q, p, t)$$

explicitly depends on time (cf. Szebehely [46, p. 55], Lichtenberg [49, p. 13]).

2.11.4   Fixed Points of a Hamiltonian System in the Plane

For the fixed points of an autonomous Hamiltonian system with one degree of freedom, we have

$$\vec{r} = \begin{pmatrix} q \\ p \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} = 0.$$

The Jacobian matrix is

$$D\vec{v} = \begin{pmatrix} \frac{\partial^2 H}{\partial p \partial q} & \frac{\partial^2 H}{\partial p^2} \\ -\frac{\partial^2 H}{\partial q^2} & -\frac{\partial^2 H}{\partial q \partial p} \end{pmatrix}.$$

If we calculate the eigenvalues of the matrix

$$A = D\vec{v}(\vec{r}_F) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.124}$$

according to Eqs. (2.107) and (2.108), we obtain

$$B = \frac{a_{11} + a_{22}}{2} = 0, \qquad C = -\det A,$$

and therefore

$$\lambda = \pm\sqrt{C} = \pm\sqrt{-\det A}.$$

Hence, the two eigenvalues are either real with opposite sign or imaginary with opposite sign.

All fixed points of the linearized system are therefore either centers or saddle points. The linearized system cannot have any sources or sinks. This is consistent with the name "conservative system."

We now regard the Hamiltonian $H(q, p)$ as a function that describes a two-dimensional surface in three-dimensional space.

The fixed-point condition

$$\begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} = 0$$

is necessary for the existence of a **relative extremum** (also called a local extremum) of $H(\vec{r})$ at $\vec{r} = \vec{r}_F$, because the gradient of H must be zero. A sufficient condition for a **relative minimum** is that the **Hessian matrix**

$$\begin{pmatrix} \frac{\partial^2 H}{\partial q^2} & \frac{\partial^2 H}{\partial q \partial p} \\ \frac{\partial^2 H}{\partial p \partial q} & \frac{\partial^2 H}{\partial p^2} \end{pmatrix}$$

of H be positive definite at $\vec{r} = \vec{r}_F$ (all eigenvalues positive). If the Hessian matrix is negative definite (all eigenvalues negative), then a **relative maximum** is present. If the Hessian matrix is indefinite (both positive and negative eigenvalues), a **saddle point** is present.

Obviously, we find for the Hessian matrix

$$\begin{pmatrix} \frac{\partial^2 H}{\partial q^2} & \frac{\partial^2 H}{\partial q \partial p} \\ \frac{\partial^2 H}{\partial p \partial q} & \frac{\partial^2 H}{\partial p^2} \end{pmatrix} = \begin{pmatrix} -a_{21} & a_{11} \\ a_{11} & a_{12} \end{pmatrix}.$$

The eigenvalues λ_H of the Hessian matrix can be determined as follows:

$$(-a_{21} - \lambda_H)(a_{12} - \lambda_H) - a_{11}^2 = 0$$

$$\Rightarrow \lambda_H^2 + \lambda_H(a_{21} - a_{12}) - (a_{11}^2 + a_{12}a_{21}) = 0$$

$$\Rightarrow \lambda_H = \frac{a_{12} - a_{21}}{2} \pm \sqrt{\frac{(a_{12} - a_{21})^2}{4} + a_{11}^2 + a_{12}a_{21}} \qquad (2.125)$$

$$\Rightarrow \lambda_H = \frac{a_{12} - a_{21}}{2} \pm \sqrt{\frac{(a_{12} + a_{21})^2}{4} + a_{11}^2}.$$

The argument of the square root is not negative, so that only real eigenvalues exist (symmetry of the Hessian matrix).

Hence, we have three possibilities for the value of the square root:

- It is greater than the absolute value of the first fraction. In this case, it determines the sign of the eigenvalues. Therefore, a positive eigenvalue and a negative eigenvalue exist, and the Hessian matrix is indefinite. Hence, we have a (geometric) saddle point. In this case, due to Eq. (2.125), we have

$$a_{11}^2 + a_{12}a_{21} > 0,$$

or with $a_{11} = -a_{22}$ (see Eq. (2.124)),

$$a_{11}a_{22} - a_{12}a_{21} < 0$$

$$\Leftrightarrow \det A < 0.$$

Due to the restriction

$$\lambda = \pm\sqrt{C} = \pm\sqrt{-\det A}$$

for the eigenvalues of the Jacobian matrix, the fixed point is also a saddle point.
- It is less than the absolute value of the first fraction in Eq. (2.125), so that $\det A > 0$ holds. Due to

$$\lambda = \pm\sqrt{-\det A},$$

the eigenvalues are imaginary, and the fixed point is a center. The first fraction in Eq. (2.125) decides which sign the eigenvalues of the Hessian matrix have. For $a_{12} > a_{21}$, we have a relative minimum of the Hamiltonian, and for $a_{12} < a_{21}$, one obtains a relative maximum.
- It equals the first fraction in Eq. (2.125). Then, one eigenvalue is zero, and $\det A = 0$ holds, which we have excluded (degenerate fixed point).

In conclusion, the Hamiltonian has a geometric saddle point if the corresponding fixed point is a saddle point. It has a relative minimum or maximum if the corresponding fixed point is a center.

2.11.5 Hamiltonian as Lyapunov Function

As in the previous sections, let us consider an autonomous Hamiltonian system with only one degree of freedom.

In Sect. 2.11.4, we saw that the linearization of such a system may have only centers and saddle points as fixed points. Let us assume that the system is linearized at a specific fixed point and that the fixed point of the linearized system is a saddle point. According to Theorem 2.17 (p. 77), the fixed point of the original system must be a saddle point as well. Theorem 2.17 applies, because the saddle point is a hyperbolic fixed point.

These arguments cannot be adopted for a center as a fixed point, because centers are not hyperbolic fixed points. If we want to show that a center of the linearized system corresponds to a center of the original nonlinear system, we need a different approach, which is presented in the following.

In many cases, the Hamiltonian of an autonomous system is defined in such a way that $H \geq 0$ holds and that for the fixed points, $H(\vec{r}_F) = 0$ is valid. If under these conditions, $H(\vec{r})$ has a minimum at $\vec{r} = \vec{r}_F$, then $L := H$ is a Lyapunov function, since one has

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \vec{v} \cdot \mathrm{grad}\, L = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} \cdot \mathrm{grad}\, H = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial H}{\partial q} \\ \frac{\partial H}{\partial p} \end{pmatrix} = 0 \leq 0.$$

Under these conditions, one is therefore able to show that the autonomous Hamiltonian system has a center.[19]

Theorem 2.19. *Let $H \in C^2(D)$ be a Hamiltonian ($D \subset \mathbb{R}^{2n}$ open). If $\vec{r}$ is an isolated minimum (strict minimum) of the Hamiltonian, then $\vec{r}$ is a stable fixed point.*

(See Amann [35, Sect. 18.11 b]; Walter [50, Sect. 30, Chap. XII d].)

Since for Hamiltonians, the question whether a minimum of maximum exists is just a matter of the sign,[20] one concludes in general the following result:

---

[19]This is not astonishing, since we obtained $\frac{\mathrm{d}L}{\mathrm{d}t} = 0$, so that L and H are constant on the trajectory.

[20]The reader may verify that if $H(q, p)$ is a Hamiltonian, then $\tilde{H} = -H$ with $\tilde{q} = p$ and $\tilde{p} = q$ is also a Hamiltonian. Furthermore, one can easily see that adding a constant to a Hamiltonian does not modify Hamilton's equations. Therefore, every relative maximum or minimum of a Hamiltonian at $\vec{r} = \vec{r}_F$ may be transformed into a relative minimum with $H(\vec{r}_F) = 0$.

Theorem 2.20. *Every nondegenerate fixed point $\vec{r}_F$ of a Hamiltonian system is a saddle point or a center. It is a saddle point if and only if the Hamiltonian has a saddle point with*

$$\det\left[D\vec{v}(\vec{r}_F)\right] < 0.$$

It is a center if and only if the Hamiltonian has a strict minimum or strict maximum with

$$\det\left[D\vec{v}(\vec{r}_F)\right] > 0.$$

(See Perko [30, Sect. 2.14, Theorem 2].)

2.11.6 Canonical Transformations

We consider **canonical transformations** as transformations that preserve the phase space area and that transform one set of Hamilton's equations (depending on q, p) into another set of Hamilton's equations (depending on Q, P).

According to Appendix A.5, preservation of the phase space area means

$$\xi = \frac{\partial(Q, P)}{\partial(q, p)} = \begin{vmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{vmatrix} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = 1. \qquad (2.126)$$

We consider only a very specific[21] subset of canonical transformations for which the value of the Hamiltonian remains unchanged. In this case,

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q},$$

---

[21]A sophisticated theory of canonical transformations exists and is described in many textbooks on theoretical physics. The transformed generalized coordinates Q_i and momenta P_i ($i \in \{1, 2, \ldots, n\}$) may depend on the original generalized coordinates q_i and momenta p_i and on the time t:

$$Q_i = Q_i(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n, t),$$
$$P_i = P_i(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n, t).$$

The Hamiltonian $H(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n, t)$ is transformed into the Hamiltonian $K(Q_1, Q_2, \ldots, Q_n, P_1, P_2, \ldots, P_n, t)$. Canonical transformations can be constructed using four basic types of **generating function** $F_1(q, Q, t)$, $F_2(q, P, t)$, $F_3(p, Q, t)$, $F_4(p, P, t)$. This theory is outside the scope of this book. An introduction may be found, for example, in [19].

must be transformed into

$$\dot{Q} = \frac{\partial H}{\partial P}, \qquad \dot{P} = -\frac{\partial H}{\partial Q}.$$

For all points in phase space we have

$$\frac{\partial H}{\partial p} = \frac{\partial H}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial p} \qquad \Rightarrow \dot{q} = -\dot{P}\frac{\partial Q}{\partial p} + \dot{Q}\frac{\partial P}{\partial p},$$

$$\frac{\partial H}{\partial q} = \frac{\partial H}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial q} \qquad \Rightarrow -\dot{p} = -\dot{P}\frac{\partial Q}{\partial q} + \dot{Q}\frac{\partial P}{\partial q}.$$

Now we have to check whether these restricted transformations are actually canonical ones, i.e., whether $\xi = 1$ holds.

For this purpose, we eliminate all derivatives of P in Eq. (2.126) by means of the last two results:

$$\xi = \frac{\partial Q}{\partial q}\left(\frac{\dot{q}}{\dot{Q}} + \frac{\dot{P}}{\dot{Q}}\frac{\partial Q}{\partial p}\right) - \frac{\partial Q}{\partial p}\left(-\frac{\dot{p}}{\dot{Q}} + \frac{\dot{P}}{\dot{Q}}\frac{\partial Q}{\partial q}\right) = \frac{1}{\dot{Q}}\left(\frac{\partial Q}{\partial q}\dot{q} + \frac{\partial Q}{\partial p}\dot{p}\right).$$

The last expression in parentheses is equal to $\dot{Q}$, so that $\xi = 1$ indeed holds.

2.11.7 Action-Angle Variables

In this section, we will briefly discuss special coordinates for oscillatory Hamiltonian systems, the so-called **action-angle variables**. Again, the general theory is outside the scope of this book, but we will use some results of this theory to determine the oscillation frequency of nonlinear systems.

2.11.7.1 Introductory Example

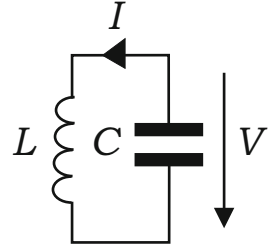As an introductory example, we consider a parallel LC circuit as shown in Fig. 2.16 for which

$$V = L\dot{I} \qquad \Rightarrow \dot{I} = \frac{V}{L}$$

and

$$-I = C\dot{V} \qquad \Rightarrow \dot{V} = -\frac{I}{C}$$

hold. For

$$q = I, \qquad p = V,$$

Fig. 2.16 Parallel LC circuit

this may be transformed into Hamilton's equations:

$$\dot{q} = \frac{p}{L} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{q}{C} = -\frac{\partial H}{\partial q}.$$

By means of an integration, we obtain the Hamiltonian

$$H = \frac{p^2}{2L} + \frac{q^2}{2C} = \frac{V^2}{2L} + \frac{I^2}{2C}.$$

For the initial conditions

$$V = V_{\max}, \qquad I = 0,$$

one obtains

$$H = \frac{V_{\max}^2}{2L}.$$

This value of the Hamiltonian is preserved, so that

$$I^2 = 2CH - \frac{C}{L}V^2 = \frac{C}{L}(V_{\max}^2 - V^2)$$

is valid. Hence, the orbit in phase space is an ellipse with semiaxes $V_{\max}$ and

$$I_{\max} = \sqrt{\frac{C}{L}}V_{\max}.$$

For the area enclosed by this orbit, one obtains

$$A = \pi V_{\max} I_{\max} = \pi \sqrt{\frac{C}{L}} V_{\max}^2 = 2\pi \sqrt{LC}\, H.$$

Since we know that the resonant angular frequency of a parallel LC circuit is

$$\omega = \omega_{\mathrm{res}} := \frac{1}{\sqrt{LC}},$$

we see at once that

$$A = \frac{2\pi}{\omega} H = TH,$$

where T is the period of the oscillation. One may therefore guess that the resonant frequency or period may be derived from the area enclosed by the orbit even if less-trivial examples are considered. If that works (and it does, as we will see soon), it will obviously not be necessary to actually solve the differential equation.

2.11.7.2   Basic Principle

Let us consider an autonomous Hamiltonian system with one degree of freedom. We assume that in the (q, p) phase space, a center exists such that closed orbits are present. We are now looking for a specific canonical transformation that introduces the new generalized coordinate/momentum pair (Q, P).

The idea of action-angle variables is to require that one of the transformed coordinates not depend on time:

$$\frac{dP}{dt} = \dot{P} = 0.$$

Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q},$$

$$\dot{Q} = \frac{\partial H}{\partial P}, \qquad \dot{P} = -\frac{\partial H}{\partial Q},$$

then show that

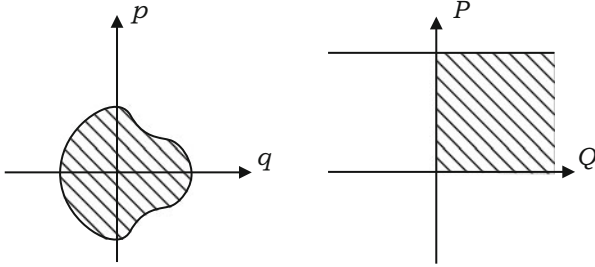$$\frac{\partial H}{\partial Q} = 0$$

is valid, so that H cannot depend on Q but only on P:

$$H = H(P).$$

Therefore, $\frac{\partial H}{\partial P}$ also depends only on P. Furthermore, P is constant with respect to time, so that

$$\frac{\partial H}{\partial P} = \frac{dH}{dP}$$

Fig. 2.17 Transition to action-angle variables

cannot depend on time either. Due to Hamilton's equation, one then obtains

$$\dot{Q} = \frac{\mathrm{d}H}{\mathrm{d}P} = K(P),$$

and therefore

$$Q(t) = Q(0) + K(P)\, t.$$

In the original phase space (q, p), one revolution lasted for time T. Hence, it is clear that in the transformed phase space (Q, P), the variable Q will increase by the amount

$$\Delta Q = K(P)\, T, \tag{2.127}$$

while P remains constant. This is visualized in Fig. 2.17.

Since the area in phase space is kept constant by a canonical transformation, the (Q, P) phase space is a surface of a cylinder (cf. Percival/Richards [51, p. 105]). This indicates why generalized coordinates are called **cyclic** if the Hamiltonian does not depend on them. In our case, Q is a cyclic coordinate.

Heretofore, we required only that P not depend on time. This is satisfied, for example, if to every point (q, p), we assign the area A that is enclosed by the orbit that goes through the point (q, p):

$$P = A = \iint_A \mathrm{d}q\, \mathrm{d}p.$$

If this definition for P is used, then shaded area in the (q, p) phase space in Fig. 2.17 is equal to P. The shaded area in the (Q, P) phase space is equal to $P\, \Delta Q$ (see the right-hand diagram in Fig. 2.17). Since both areas must be equal, we obtain

$$\Delta Q = 1,$$

and therefore, based on Eq. (2.127),

$$K(P)\, T = 1$$

$$\Leftrightarrow K(P) = \frac{\mathrm{d}H}{\mathrm{d}P} = \frac{1}{T}.$$

In conclusion, we may use H to calculate the period of the oscillation directly without solving the differential equation explicitly.

Instead of taking the area $A = P$ directly as a generalized coordinate, one defines the **action variable**

$$J = \frac{P}{2\pi}$$

$$\Leftrightarrow \boxed{J = \frac{1}{2\pi} \iint_A \mathrm{d}q\, \mathrm{d}p}$$

and the **angle variable**[22]

$$\theta = 2\pi\, Q.$$

As the name implies, the angle variable obviously increases by 2π during every period of the oscillation. Hence, one obtains

$$\frac{\mathrm{d}H}{\mathrm{d}J} = \frac{\mathrm{d}H}{\mathrm{d}P}\, 2\pi = \frac{2\pi}{T}$$

$$\Rightarrow \boxed{\frac{\mathrm{d}H}{\mathrm{d}J} = \frac{2\pi}{T} = \omega}$$

for Hamilton's equations

$$\boxed{\dot\theta = \frac{\partial H}{\partial J}, \qquad \dot J = -\frac{\partial H}{\partial \theta} = 0.}$$

Please note that for these considerations, we assumed that the Hamiltonian does not depend on time and that the orbits are closed. Therefore, the action variable is

---

[22]It is easy to show that multiplying the generalized coordinate by a constant and dividing the generalized momentum by that same constant is a canonical transformation (cf. Sect. 5.1.3).

defined in a unique way, and by Liouville's theorem, it is obvious that the phase space area, and thus also the action variable, remains constant.

2.11.8　LC Circuit with Nonlinear Inductance

The characteristic curve $B(H)$ of a magnetic material can be approximated by

$$B = B_{\text{max}} \frac{2}{\pi} \arctan \frac{H}{H_0} = B_{\text{max}} \frac{2}{\pi} \arctan \frac{I}{I_0}.$$

The magnetic material will be used to build an inductor with N windings. With the magnetic flux $\Phi_{\text{m}} = BA$, it follows that

$$V = N \frac{d\Phi_{\text{m}}}{dt} = N \, A B_{\text{max}} \frac{2}{\pi} \frac{1}{1 + \left(\frac{I}{I_0}\right)^2} \frac{1}{I_0} \frac{dI}{dt}.$$

Therefore, from

$$V = L \frac{dI}{dt},$$

one obtains

$$L(I) = \frac{L_0}{1 + \left(\frac{I}{I_0}\right)^2}.$$

The corresponding inductor in parallel with a capacitor can now be used to form an LC oscillator as shown in Fig. 2.16. For the capacitance of the oscillating circuit,

$$-I = C \frac{dV}{dt}.$$

is valid. The magnetic energy is

$$W_{\text{magn}} = \int V \, I \, dt = \int L(I) \, I \, \frac{dI}{dt} \, dt = L_0 \int \frac{I}{1 + \left(\frac{I}{I_0}\right)^2} \frac{dI}{dt} \, dt$$

$$= L_0 I_0^2 \int \frac{x}{1 + x^2} \frac{dx}{dt} \, dt,$$

where we have used

$$x = \frac{I}{I_0}, \quad \text{or} \quad I = x \, I_0.$$

Because of

$$\int \frac{x}{1+x^2}\, dx = \frac{1}{2}\, \ln|1+x^2| + \text{const},$$

one obtains

$$W_{\text{magn}} = \frac{L_0 I_0^2}{2}\, \ln\left(1 + \left(\frac{I}{I_0}\right)^2\right).$$

Together with

$$W_{\text{el}} = \frac{1}{2} C V^2,$$

this leads to

$$W(V, I) = W_{\text{el}} + W_{\text{magn}} = \frac{1}{2} C V^2 + \frac{L_0 I_0^2}{2}\, \ln\left(1 + \left(\frac{I}{I_0}\right)^2\right).$$

If we define $p = V$, we obtain

$$\frac{\partial W}{\partial p} = Cp = CV = CL(I)\frac{dI}{dt} = \frac{CL_0}{1 + \left(\frac{I}{I_0}\right)^2}\frac{dI}{dt}.$$

If this is one of the two Hamilton's equations, the right-hand side must be equal to $\dot{q}$, and one obtains

$$q = CL_0 I_0\ \arctan\frac{I}{I_0}.$$

Therefore, the Hamiltonian is

$$H(q, p) = \frac{1}{2} C p^2 + \frac{L_0 I_0^2}{2}\, \ln\left(1 + \tan^2\frac{q}{CL_0 I_0}\right)$$

$$\Rightarrow H(q, p) = \frac{1}{2} C p^2 - L_0 I_0^2\, \ln\, \cos\frac{q}{CL_0 I_0}.$$

We still have to check the second of Hamilton's equations. One obtains

$$\frac{\partial H}{\partial q} = -L_0 I_0^2\, \frac{1}{\cos\frac{q}{CL_0 I_0}}\left(-\sin\frac{q}{CL_0 I_0}\right)\frac{1}{CL_0 I_0} = \frac{I_0}{C}\, \tan\frac{q}{CL_0 I_0} = \frac{I}{C}.$$

In fact, the right-hand side equals $-\frac{dV}{dt}$, which is equal to $-\dot{p}$, and both Hamilton's equations are satisfied.

In order to calculate the oscillation frequency, we compute the action:

$$
\begin{aligned}
J(H) &= \frac{1}{2\pi} \iint dq \, dp \\
&= \frac{1}{\pi} \int_{q_1}^{q_2} p \, dq \\
&= \frac{1}{\pi} \sqrt{\frac{2}{C}} \int_{q_1}^{q_2} \sqrt{H + L_0 I_0^2 \, \ln \, \cos \frac{q}{C L_0 I_0}} \, dq.
\end{aligned}
$$

The limits q_1 and q_2 are determined by the zeros of p where the trajectory crosses the q-axis. The substitution

$$
x = \frac{q}{C L_0 I_0}, \qquad \frac{dx}{dq} = \frac{1}{C L_0 I_0},
$$

leads to

$$
J(H) = \frac{1}{\pi} \sqrt{\frac{2}{C}} C L_0 I_0 \int_{x_1}^{x_2} \sqrt{H + L_0 I_0^2 \, \ln \, \cos x} \, dx. \qquad (2.128)
$$

Due to

$$
\ln \, \cos x \approx -\frac{x^2}{2} - \frac{x^4}{12} - \cdots
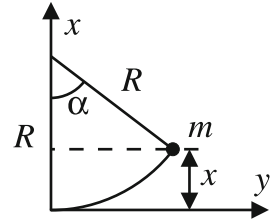$$

the simplest approximation for $I \ll I_0$ is

$$
\begin{aligned}
J(H) &= \frac{1}{\pi} \sqrt{2C} L_0 I_0 \sqrt{\frac{L_0}{2}} I_0 \int_{x_1}^{x_2} \sqrt{H \frac{2}{L_0 I_0^2} - x^2} \, dx \\
&= \frac{1}{\pi} \sqrt{L_0 C} L_0 I_0^2 \int_{x_1}^{x_2} \sqrt{H \frac{2}{L_0 I_0^2} - x^2} \, dx.
\end{aligned}
$$

The integral describes the area of a semicircle with radius

$$
\sqrt{\frac{2H}{L_0 I_0^2}},
$$

Fig. 2.18 Mathematical pendulum

so that

$$J(H) = \frac{1}{\pi} \sqrt{L_0 C} \; L_0 I_0^2 \; \frac{2H}{L_0 I_0^2} \frac{\pi}{2} = \sqrt{L_0 C} \; H$$

is obtained. As expected, one obtains

$$\omega = \frac{\mathrm{d}H}{\mathrm{d}J} = \frac{1}{\sqrt{L_0 C}}.$$

If the approximation is undesirable, one may directly calculate the derivative of Eq. (2.128). Then the integral may be evaluated numerically in order to calculate the amplitude-dependent oscillation frequency $\omega(\hat{I})$. As mentioned above, no direct solution of the differential equation is required.

2.11.9   Mathematical Pendulum

Consider a mathematical pendulum with mass m that is suspended by means of a massless cord of length R (see Fig. 2.18). Suppose that initially, the mass m is at height $x = h$ (corresponding to the angle $\alpha = \hat{\alpha}$) with zero velocity.

2.11.9.1   Energy Balance

The sum of the potential energy and kinetic energy must remain constant:

$$W_{\mathrm{pot}} + W_{\mathrm{kin}} = \mathrm{const}, \qquad \text{with} \quad W_{\mathrm{pot}} = mgx \qquad \text{and} \quad W_{\mathrm{kin}} = \frac{1}{2}mu^2$$

$$\Leftrightarrow mgx + \frac{1}{2}mu^2 = \mathrm{const}$$

$$\Leftrightarrow g(R - R \, \cos\alpha) + \frac{1}{2}R^2\dot{\alpha}^2 = \mathrm{const}.$$

We now calculate the time derivative of this equation:

$$gR \, \sin\alpha \, \dot\alpha + R^2 \dot\alpha \ddot\alpha = 0.$$

As a result, we obtain

$$\ddot\alpha + \frac{g}{R} \, \sin\alpha = 0. \tag{2.129}$$

2.11.9.2   Hamilton's Equations

We now try to convert Eq. (2.129) into a pair of Hamilton's equations using our standard approach

$$q = \alpha, \qquad p = \dot\alpha,$$

which leads to

$$\dot q = p,$$
$$\dot p = -\frac{g}{R} \, \sin q.$$

If this is to be in accord with Hamilton's equations,

$$\frac{\partial H}{\partial p} = \dot q,$$
$$\frac{\partial H}{\partial q} = -\dot p,$$

we obtain by integration

$$H = \frac{p^2}{2} + f(q),$$
$$H = -\frac{g}{R} \, \cos q + g(p).$$

Putting both results together, one obtains

$$H(q, p) = \frac{p^2}{2} - \frac{g}{R} \, \cos q. \tag{2.130}$$

If we additionally require $H(0, 0) = 0$, we may add a constant accordingly:

$$H(q, p) = \frac{p^2}{2} + \frac{g}{R}\, (1 - \cos q). \tag{2.131}$$

2.11.9.3   Oscillation Period

In order to calculate the oscillation period, we first determine the action variable:

$$J(H) = \frac{1}{2\pi} \iint \mathrm{d}q\, \mathrm{d}p = \frac{1}{\pi} \int_{q_1}^{q_2} p\, \mathrm{d}q.$$

Equation (2.130) leads to

$$p = \sqrt{2\left[H + \frac{g}{R}\, \cos q\right]}$$

for the upper part of the curve in phase space. By means of

$$a = \frac{2}{\pi^2} H,$$

$$b = \frac{2}{\pi^2} \frac{g}{R}, \tag{2.132}$$

$$\frac{a}{b} = \frac{HR}{g} \tag{2.133}$$

one obtains the following integral:

$$J(H) = \int_{q_1}^{q_2} \sqrt{a + b\, \cos q}\, \mathrm{d}q. \tag{2.134}$$

The limits q_1 and q_2 of integration are determined by the zeros of p. We obviously have $p = 0$ for

$$q_{1,2} = \mp \arccos \frac{-a}{b}.$$

Therefore, $q_1 = -q_2$ holds, so that we can use the symmetry of the integrand:

$$J(H) = 2 \int_0^{q_2} \sqrt{a + b\, \cos q}\, \mathrm{d}q.$$

According to the first formula 2.576 in [3], for $|a| \le b$ and $0 \le q < \arccos(-a/b)$, the integral has the following value:[23]

$$J(H) = 2\sqrt{\frac{2}{b}} \left[(a-b) \, \text{F}\left(\gamma, \frac{1}{r}\right) + 2b \, \text{E}\left(\gamma, \frac{1}{r}\right) \right]_0^{q_2} \qquad (2.135)$$

with

$$r = \sqrt{\frac{2b}{a+b}} \qquad \text{and} \qquad \gamma = \arcsin \sqrt{\frac{b(1-\cos q)}{a+b}}.$$

For $q = 0$, one obviously has $\gamma = 0$. For $q = q_2$,

$$\cos q = \frac{-a}{b} \qquad \Rightarrow \gamma = \arcsin \sqrt{\frac{b\left(1 + \frac{a}{b}\right)}{a+b}} = \arcsin 1 = \frac{\pi}{2}$$

is valid. The expression in square brackets in Eq. (2.135) is equal to zero at the lower integration limit $\gamma = 0$, since $\text{F}(0, k) = 0$ and $\text{E}(0, k) = 0$:

$$(a-b)\text{F}(0, k) + 2b \, \text{E}(0, k) = 0.$$

Here we set

$$k = \frac{1}{r} = \sqrt{\frac{1}{2} + \frac{a}{2b}} \qquad \text{and} \qquad k' = \sqrt{1 - k^2} = \sqrt{\frac{1}{2} - \frac{a}{2b}}. \qquad (2.136)$$

From $\text{F}(\pi/2, k) = \text{K}(k)$ and $\text{E}(\pi/2, k) = \text{E}(k)$, one concludes, based on Eq. (2.135), that

$$J(H) = 2\sqrt{\frac{2}{b}} \left[(a-b) \, \text{K}(k) + 2b \, \text{E}(k) \right] = 4\sqrt{2b} \left[\text{E}(k) - k'^2 \text{K}(k) \right]. \qquad (2.137)$$

---

[23] Here $\text{F}(\gamma, k)$ denotes the **elliptic integral of the first kind**, whereas $\text{E}(\gamma, k)$ is an **elliptic integral of the second kind** [52]:

$$\text{F}(\gamma, k) = \int_0^\gamma \frac{d\theta}{\sqrt{1 - k^2 \, \sin^2 \theta}} \qquad \text{E}(\gamma, k) = \int_0^\gamma \sqrt{1 - k^2 \, \sin^2 \theta} \; d\theta.$$

The **complete elliptic integral of the first kind** is defined by $\text{K}(k) = \text{F}(\pi/2, k)$, while the **complete elliptic integral of the second kind** is given by $\text{E}(k) = \text{E}(\pi/2, k)$. In this book, we make use of only the **modulus** k ($0 \le k \le 1$). Alternatively, one can also use the **parameter** m or the **modular angle** α:

$$k = \sin \alpha, \qquad m = k^2.$$

The **complementary modulus** k' is given by $k^2 + k'^2 = 1$, and the **complementary parameter** $m_1 = k'^2$ is therefore defined by $m + m_1 = 1$.

The angular frequency of the oscillation may be calculated according to

$$\omega = \frac{\partial H}{\partial J}.$$

Since by definition, H depends only on the action variable J but not on the angle variable θ, and since H does not depend on time in our case, the partial derivative is in fact a total derivative:

$$\omega = \frac{\partial H}{\partial J} = \frac{dH}{dJ} = \left(\frac{dJ}{dH}\right)^{-1}.$$

We therefore need $\frac{dJ}{dH}$. From Eqs. (2.133), (2.136), and (2.137), we get

$$\frac{dJ}{dH} = \frac{dJ}{dk}\frac{dk}{dH} = 4\sqrt{2b}\left[\frac{dE(k)}{dk} - 2k'\frac{-2k}{2k'}K(k) - (1-k^2)\frac{dK(k)}{dk}\right]\frac{1}{2k}\frac{R}{2g} =$$

$$= \frac{2}{\pi k}\sqrt{\frac{R}{g}}\left[\frac{dE(k)}{dk} + 2k\,K(k) - (1-k^2)\frac{dK(k)}{dk}\right].$$

In the last step, we made use of Eq. (2.132), which led to

$$\sqrt{2b} = \frac{2}{\pi}\sqrt{\frac{g}{R}}.$$

With

$$\boxed{\frac{dK(k)}{dk} = \frac{E(k)}{kk'^2} - \frac{K(k)}{k}}$$

and

$$\boxed{\frac{dE(k)}{dk} = \frac{E(k) - K(k)}{k},}$$

we obtain

$$\frac{dJ}{dH} = \frac{2}{\pi k}\sqrt{\frac{R}{g}}\frac{E(k) - K(k) + 2k^2\,K(k) - E(k) + k'^2 K(k)}{k} = \frac{2}{\pi}\sqrt{\frac{R}{g}}\,K(k).$$

This finally leads to

$$\omega = \frac{dH}{dJ} = \frac{\pi}{2K(k)}\sqrt{\frac{g}{R}}. \tag{2.138}$$

The calculation presented here can be simplified significantly if the derivative with respect to H is determined before the integral is evaluated. This is done in Sect. 3.16 for an analogous problem.

Now our considerations are complete in principle. Only the geometric meaning of the modulus k remains to be clarified.

From Eq. (2.136), one obtains

$$k = \sqrt{\frac{1}{2} + \frac{HR}{2g}}.$$

Initially, the mass is momentarily at rest, so that we have $p = \dot{\alpha} = 0$. Therefore, according to Eq. (2.130),

$$H = -\frac{g}{R}\cos\hat{\alpha}$$

is the value of the Hamiltonian (which remains constant). This leads to

$$k = \sqrt{\frac{1 - \cos\hat{\alpha}}{2}}.$$

Since

$$\cos\hat{\alpha} = \cos^2\frac{\hat{\alpha}}{2} - \sin^2\frac{\hat{\alpha}}{2} = 1 - 2\,\sin^2\frac{\hat{\alpha}}{2},$$

this may be written in the form

$$k = \sin\frac{\hat{\alpha}}{2}.$$

2.11.10 Vlasov Equation

From the formula

$$\operatorname{div}(\rho\,\vec{v}) = \vec{v}\cdot\operatorname{grad}\rho + \rho\operatorname{div}\vec{v},$$

which is known from vector analysis, the continuity equation (2.115) for incompressible flows leads to the differential equation

$$-\dot{\rho} = \vec{v}\cdot\operatorname{grad}\rho.$$

If we consider a Hamiltonian system with one degree of freedom that describes the incompressible flow, we have

$$\vec{r} = \begin{pmatrix} q \\ p \end{pmatrix}, \qquad \vec{v} = \dot{\vec{r}} = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix}.$$

Therefore, one obtains

$$\frac{\partial \rho}{\partial t} + \dot{q}\,\frac{\partial \rho}{\partial q} + \dot{p}\,\frac{\partial \rho}{\partial p} = 0$$

$$\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial H}{\partial p}\,\frac{\partial \rho}{\partial q} - \frac{\partial H}{\partial q}\,\frac{\partial \rho}{\partial p} = 0.$$

This is the **Vlasov equation**. It describes how the particle density ρ at different locations changes with time.

2.11.11 Outlook

A dynamical system is called **conservative** if the total energy (or the area in phase space) remains constant.

Every autonomous Hamiltonian system is conservative. However, there exist non-Hamiltonian systems that are conservative.

A function $I(q_k, p_k)$ that does not depend on t and that does not change its value on the trajectory is called a **constant of the motion**. Such a constant of the motion allows one to reduce the order of the problem by 1 (cf. Tabor [53, p. 2]), since one may express one variable in terms of the other variables by means of this function.

For a non-Hamiltonian system of order n, one therefore needs $n - 1$ constants of the motion in order to completely solve the differential equation by means of quadratures (cf. Tabor [53, p. 39]).

A Hamiltonian system is called **integrable** if the solution can be determined by quadratures (cf. Rebhan [22, vol. I, p. 287]). This is the case if the problem can be written in action-angle variables.

In contrast to non-Hamiltonian systems, one needs only n constants of the motion if a Hamiltonian system of order $2n$ with n degrees of freedom is considered (instead of $2n - 1$, as in the general case).

Conservative Hamiltonian systems with one degree of freedom (order 2) are integrable (cf. Rebhan [22, vol. I, p. 359]). This is obvious, because the Hamiltonian itself is a constant of the motion.

Chaotic behavior is possible only in nonintegrable systems (cf. Rebhan [22, vol. I, pp. 335 and 359]). Therefore, chaos is not possible in autonomous Hamiltonian systems with one degree of freedom. However, if more degrees of freedom are present, chaotic behavior may also occur in autonomous Hamiltonian systems.

References

1. A.H. Zemanian, *Distribution Theory and Transform Analysis* (Dover, New York, 1987)
2. K.B. Howell, *Principles of Fourier Analysis* (Chapman & Hall/CRC, Boca Raton/London/New York/Washington, 2001)
3. I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, 6th edn. (Academic, San Diego/San Francisco/New York/Boston/London/Sydney/Tokyo, 2000)
4. D. Müller-Wichards, *Transformationen und Signale* (B. G. Teubner, Stuttgart/Leipzig, 1999)
5. B. Davies, *Integral Transforms and Their Applications*, 3rd edn. (Springer, Berlin/Heidelberg/New York, 2002)
6. W.R. LePage, *Complex Variables and the Laplace Transform for Engineers* (McGraw-Hill, New York/Toronto/London, 1961)
7. B. Girod, R. Rabenstein, A. Stenger, *Einführung in die Systemtheorie*, 2. Auflage (B. G. Teubner, Stuttgart/Leipzig/Wiesbaden, 2003)
8. E. Kreyszig, *Statistische Methoden und ihre Anwendungen* (Vandenhoeck & Ruprecht, Göttingen, 1967)
9. A.W. Chao, M. Tigner, *Handbook of Accelerator Physics and Engineering*, 3rd edn. (World Scientific/New Jersey/London/Singapore/Beijing/Shanghai/Hong Kong/Taipei/Chennai, 2006)
10. M. Reiser, *Theory and Design of Charged Particle Beams* (Wiley, New York/Chichester/Brisbane/Toronto/Singapore, 1994)
11. J.A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York/London, 1941)
12. R.E. Collin, *Field Theory of Guided Waves* (IEEE Press, New York, 1990)
13. J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (Wiley, New York/Chichester/Weinheim/Brisbane/Singapore/Toronto, 1998)
14. J.D. Jackson, *Klassische Elektrodynamik*, 3. Auflage (Walter de Gruyter, Berlin/New York, 2002)
15. G. Lehner, *Electromagnetic Field Theory for Engineers and Physicists* (Springer, Heidelberg/Dordrecht/London/New York, 2010)
16. G. Lehner, *Elektromagnetische Feldtheorie für Ingenieure und Physiker*, 3. Auflage (Springer, Berlin/Heidelberg/New York/Barcelona/Budapest/Hongkong/London/Mailand/Paris/Santa Clara/Singapur/Tokio, 1996)
17. H. Klingbeil, *Elektromagnetische Feldtheorie. Ein Lehr- und Übungsbuch*, 2. Auflage (B. G. Teubner, Stuttgart/Leipzig/Wiesbaden, 2010)
18. W. Rindler, *Introduction to Special Relativity* (Clarendon Press, Oxford, 1982)
19. H. Goldstein, C. Poole, J. Safko, *Classical Mechanics*, 3rd edn. (Pearson Education, Addison-Wesley, June 25, 2001). ISBN-10: 0201657023, ISBN-13: 978-0201657029
20. C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation* (W. H. Freeman and Company, San Francisco, 1973)
21. E. Schmutzer, *Grundlagen der Theoretischen Physik* (Deutscher Verlag der Wissenschaften, Berlin, 1989)
22. E. Rebhan, *Theoretische Physik*, Band 1, 1. Auflage (Spektrum Akademischer Verlag, Heidelberg/Berlin, 1999)
23. H. Goenner, *Einführung in die spezielle und allgemeine Relativitätstheorie*, 1. Auflage (Spektrum Akademischer Verlag, Heidelberg/Berlin/Oxford, 1996)
24. T. Fließbach, *Elektrodynamik* (Bibliographisches Institut & F. A. Brockhaus AG, Mannheim/Leipzig/Wien/Zürich, 1994)
25. T. Fließbach, *Allgemeine Relativitätstheorie*, 4. Auflage (Spektrum Akademischer Verlag, Heidelberg/Berlin, 2003)
26. W. Greiner, J. Rafelski, *Theoretische Physik, Band 3A: Spezielle Relativitätstheorie*, 2. Auflage (Verlag Harri Deutsch, Frankfurt am Main, 1989)
27. H. Mitter, *Mechanik* (Bibliographisches Institut & F. A. Brockhaus AG, Mannheim/Wien/Zürich, 1989)

28. H. Mitter, *Elektrodynamik*, 2. Auflage (Bibliographisches Institut & F. A. Brockhaus AG, Mannheim/Wien/Zürich, 1990)
29. B. Aulbach, *Gewöhnliche Differenzialgleichungen*, 2. Auflage (Spektrum, München, 2004)
30. L. Perko, *Differential Equations and Dynamical Systems* (Springer, New York/Berlin/ Heidelberg, 2002)
31. G.J. Wirsching, *Gewöhnliche Differentialgleichungen* (B. G. Teubner Verlag/GWV Fachverlage GmbH, Wiesbaden, 2006)
32. V.I. Arnold, *Gewöhnliche Differentialgleichungen*, 2. Auflage (Springer, Berlin/Heidelberg/ New York, 2001)
33. J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York/Berlin/Heidelberg/Tokyo, 1986)
34. E. Zehnder, *Lectures on Dynamical Systems. Hamiltonian Vector Fields and Symplectic Capacities* (European Mathematical Society, Zürich, 2010)
35. H. Amann, *Gewöhnliche Differentialgleichungen*, 2. Auflage (de Gruyter, Berlin/New York, 1995)
36. S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*, 1st edn. (Springer, New York /Berlin/Heidelberg/Barcelona/Hong Kong/London/Milan/Paris/Singapore/Tokyo, 1999)
37. J. La Salle, S. Lefschetz, *Die Stabilitätstheorie von Ljapunov* (Bibliographisches Institut AG, Mannheim, 1967)
38. S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2nd edn. (Springer, New York/Berlin/Heidelberg/Hong Kong/London/Milan/Paris/Tokyo, 2003)
39. H. Amann, *Ordinary Differential Equations. An Introduction to Nonlinear Analysis* (Walter de Gruyter & Co., Berlin/New York, 1990)
40. K. Burg, H. Haf, F. Wille, *Höhere Mathematik für Ingenieure* (B. G. Teubner, Stuttgart, 2002) (Bd. I: 5. Aufl. 2001, Bd. II: 4. Aufl. 2002, Bd. III: 4. Aufl. 2002, Bd. IV: 2. Aufl. 1994, Bd. V: 2. Aufl. 1993)
41. E.A. Jackson, *Perspectives of Nonlinear Dynamics* (Cambridge University Press, Cambridge/New York/New Rochelle/Melbourne/Sydney, 1991)
42. W.E. Boyce, R.C. DiPrima, *Gewöhnliche Differentialgleichungen*, 1. Auflage (Spektrum Akademischer Verlag, Heidelberg/Berlin/Oxford, 1995)
43. W.E. Boyce, R.C. DiPrima, *Elementary Differential Equations*, 5th edn. (Wiley, New York, 1991)
44. I.N. Bronstein, K.A. Semendjajew, *Taschenbuch der Mathematik*, 24. Auflage (Verlag Harri Deutsch, Thun und Frankfurt am Main, 1989)
45. U. van Rienen, *Numerical Methods in Computational Electrodynamics. Linear Systems in Practical Applications* (Springer, Berlin/Heidelberg/New York/Barcelona/Hong Kong/ London/Milan/Paris/Singapore/Tokyo, 2001)
46. V. Szebehely, *Theory of Orbits* (Academic, New York/San Francisco/London, 1967)
47. A.L. Fetter, J.D. Walecka, *Theoretical Mechanics of Particles and Continua* (McGraw-Hill, New York/St. Louis/San Francisco/Auckland/Bogotá/Hamburg/Johannesburg/London/Madrid/ Mexico/Montreal/New Delhi/Panama/Paris/São Paulo/Singapore/Sydney/Tokyo/Toronto, 1980)
48. A. Budó, *Theoretische Mechanik*, 9. Auflage (Deutscher Verlag der Wissenschaften, Berlin, 1978)
49. A.J. Lichtenberg, M.A. Liebermann, *Regular and Stochastic Motion* (Springer, New York/ Heidelberg/Berlin, 1983)
50. W. Walter, *Gewöhnliche Differentialgleichungen*, 7. Auflage (Springer, Berlin/Heidelberg/ New York, 2000)
51. I. Percival, D. Richards, *Introduction to Dynamics* (Cambridge University Press, Cambridge/ New York/Melbourne, 1982)
52. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965)
53. M. Tabor, *Chaos and Integrability in Nonlinear Dynamics* (Wiley, New York/Chichester/ Brisbane/Toronto/Singapore, 1989)