1

Mechanics

We start with an outline of classical mechanics, to provide a framework for the discrete element method (DEM). While most of the material in this chapter can be found scattered in various books on mechanics, no text seems to be available which covers concisely the concepts needed for DEM simulation. This chapter is intended as a crash course in theoretical mechanics, with an emphasis on issues relevant to computer implementation and testing. We give a list of secondary literature that the reader may refer to for further details.
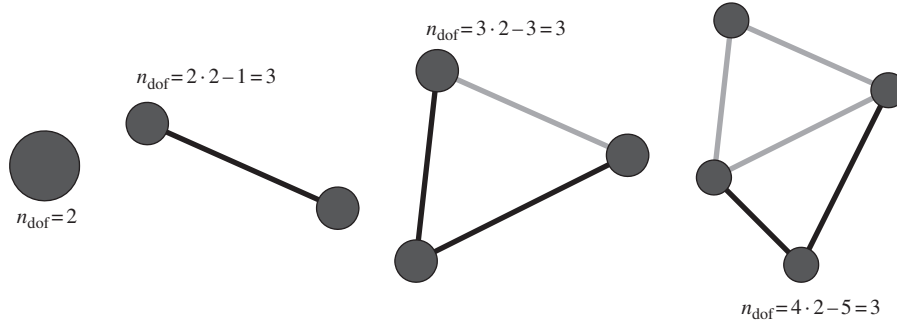
1.1 Degrees of freedom

Before discussing the dynamics of a mechanical system, we need to understand the nature of the variables in the system. There are independent variables on the one hand, usually called 'degrees of freedom', and then there are dependent variables which depend on the degrees of freedom, via algebraic relations or derivatives.

1.1.1 Particle mechanics and constraints

The concept of a 'mass point' means that we neglect the size of the mass and are interested only in its trajectory. The position of a single mass point moving along the Cartesian x-axis is described by the value of x, which corresponds to a single degree of freedom. A point moving in the xy-plane has two degrees of freedom, $\mathbf{r}_{2D} = (x, y)$, and a point moving in three-dimensional real space will have three degrees of freedom, $\mathbf{r}_{3D} = (x, y, z)$. Although we can describe the motion of a point in three-dimensional space by four 'space–time coordinates' using the tuple (x, y, z, t), in classical mechanics t is not considered a degree of freedom but rather a parameter, i.e. an independent variable which cannot be influenced.

Two mass points moving independently along the x-axis represent two degrees of freedom, $\mathbf{r}_1$ and $\mathbf{r}_2$ (here and in the following, we assume equal masses). If we 'glue' these two particles together at distance $d = \mathbf{r}_1 - \mathbf{r}_2$ as in Figure 1.1, one degree of freedom gets lost, and we are

Figure 1.1 In two dimensions, the number of degrees of freedom n_{dof} for 1, 2, 3 or 4 constrained particles with an increasing number of constraints introduced. Newly added constraints are in black; previous constraints are in gray.

left with only a single degree of freedom; in this case we can use either of $\mathbf{r}_1$, $\mathbf{r}_2$ or the average $(\mathbf{r}_1 + \mathbf{r}_2)/2$ to determine the position of both particles uniquely. This means that one constraint between two position variables eliminates one degree of freedom.

In two dimensions, for two point particles at $\mathbf{r}_1 = (x_1, y_1)$ and $\mathbf{r}_2 = (x_2, y_2)$ we have four degrees of freedom, x_1, y_1, x_2 and y_2. If we again fix the distance between the particles at a constant distance d, so that

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = d, \tag{1.1}$$

we can choose any three variables from $\{x_1, y_1, x_2, y_2\}$ and the fourth will then be determined from (1.1) by elementary geometry. Alternatively, we can introduce new variables, such as the position of the center of mass, $(x, y) = (\mathbf{r}_1 + \mathbf{r}_2)/2$ for particles of the same mass, the displacement $(x, y) = (x_2 - x_1, y_2 - y_1)$ between the particles, and the angle θ that the line segment between the two particles makes with the x-axis. In any case, we end up with three independent variables to describe the positions of the two particles fully. This means that a single constraint (1.1) reduces the number of degrees of freedom, i.e. the number of independent variables in the system, by 1.

In three-dimensional space, for two particles at positions (x_1, y_1, z_1) and (x_2, y_2, z_2) as shown in Figure 1.2, a constraint

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} = d \tag{1.2}$$

will again reduce the number of degrees of freedom by 1, so if we want to work with the center of mass

$$(x, y, z) = \frac{1}{2}\{(x_1, y_1, z_1) + (x_2, y_2, z_2)\},$$

we need two angles, ϕ and θ say, to describe the orientation of the 'rod' in space. Rotation around the orientation of the rod is not a degree of freedom, as it does not change the positions of the two points. In principle, it does not matter how one defines the degrees of freedom, whether it is with six variables and one constraint (1.2), with three Cartesian coordinates for
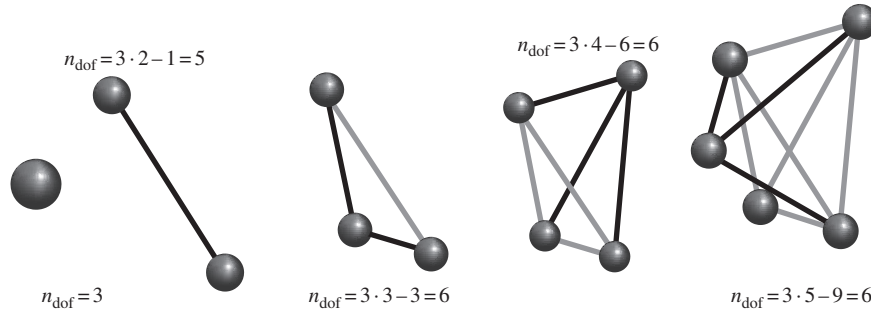
Figure 1.2 In three dimensions, the number of degrees of freedom n_{dof} for 1, 2, 3, 4 or 5 particles constrained so that the resulting cluster has no internal degrees of freedom. Newly added constraints are in black; previous constraints are in gray.

the center of mass and two angles, or with three Cartesian coordinates for one endpoint and two angles. In each case the number of degrees of freedom is the same, namely 5.

1.1.2 From point particles to rigid bodies

When we introduce one more point mass at (x_3, y_3, z_3) to our set-up, we have 9 variables in total. If we connect this new point to both ends of our rod with the additional constraints

$$\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2 + (z_3 - z_1)^2} = d_2, \tag{1.3}$$

$$\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2 + (z_3 - z_2)^2} = d_3, \tag{1.4}$$

we get a triangle, as in the middle diagram of Figure 1.2. Again, we can give an alternative description of its position in space using the center of mass, and use three angles, ϕ, θ and ψ, to describe the orientation. So the formula

$$\underbrace{(\text{degrees of freedom})}_{6} = \underbrace{(\text{variables})}_{9} - \underbrace{(\text{constraints})}_{3}$$

again holds. If we connect a fourth particle rigidly to the cluster of three particles so that it does not lie in the plane described by the other three, as shown in the fourth diagram from the left in Figure 1.2, then the three extra constraints exactly compensate for the additional three coordinates (x_4, y_4, z_4) of the new particle. In fact, for four or more spatially connected particles, the total number of degrees of freedom is always 6. Note that the rigid body formed by the connected particles need not be three-dimensional; for example, although a triangle is a two-dimensional shape, if it can rotate in three dimensions, then it also has six degrees of freedom. Through the reasoning above, we have derived that an extended rigid body has six degrees of freedom, irrespective of its size. The angular degrees of freedom ϕ, θ, ψ are obtained from the rectilinear degrees of freedom (x_1, y_1, z_1), (x_2, y_2, z_2), ... of the particles upon introducing constraints of finite length between the particles.

'Mathematically' one can define a point particle as an object having 'zero extension' and a rigid body as one having 'zero deformation'. A more pragmatic definition of a point particle is an object whose extent is much smaller than the distances that it covers in the processes under investigation; after all, the Earth is pretty extended, but the point-mass approach to describing its trajectory around the sun works rather well. Likewise, a rigid body is an object for which the deformations are much smaller than the scales that are of interest in the processes being investigated.

### 1.1.3   More context and terminology

In principle, a 'continuum' has infinitely many degrees of freedom; but in order to solve continuum problems with a computer, we have to first discretize the continuum to obtain a finite number of degrees of freedom. We could, for instance, decompose the continuum into representative mass points and model the elasticity by springs between the mass points. The deformation of a spring can be computed from the positions of the bodies, so the springs will not be degrees of freedom, while the coordinates of the mass points will be degrees of freedom. With a finite element discretization, we decompose the elastic continuum into a space-filling partition of elements for which elastic stress relations hold, and the degrees of freedom are the nodes of the elements. Depending on the choice of boundary conditions, there may be as many nodes as there are elements, or more; therefore, from the nodes one can calculate the center of mass of the elements, but not vice versa. Describing the physics via the motion of particles, for example of centers of mass, is called the 'Lagrangian representation'. This approach is natural for particulate systems, so we will adopt it in this book. Formulating the physics for a reference system in which, e.g., density amplitudes change is called the 'Eulerian representation'; this representation is preferable for many continuum problems. In a Lagrangian representation, velocities of mechanical bodies are not degrees of freedom: they can be obtained as the time derivatives of the positions on which they depend. On the other hand, when we simulate a fluid volume where velocities are assigned to the nodes of a finite element or finite difference approximation in 'Eulerian representation', it is the velocities that are the degrees of freedom.

In the previous two subsections, we introduced constraints as algebraic relations between positions, but we remark here that constraints (whose associated functions are usually denoted by g in formulae) can also be imposed on velocities. For a pendulum of length l swinging around the origin as in Figure 1.3(a), the constraint $g(x, y)$ stating that the bob (whose diameter we will neglect) stays at constant distance from the origin is

$$x^2 + y^2 = l^2. \tag{1.5}$$

In § 2.8 we will discuss the numerical solution of a problem where, in addition to constraints on x and y, constraint relations for $\dot{x}$ and $\dot{y}$ are also in effect. In undergraduate mechanics, it is common to circumvent solving the equations of motion of a constrained system with variables (x, y) that simultaneously satisfy (1.5) by transforming into plane polar coordinates (ϕ, r) so that r is eliminated. For more complicated mechanical systems, such a simplifying transformation may not be possible any more, for instance if the pendulum is connected with a unidirectionally moving body as in Figure 1.3(b).
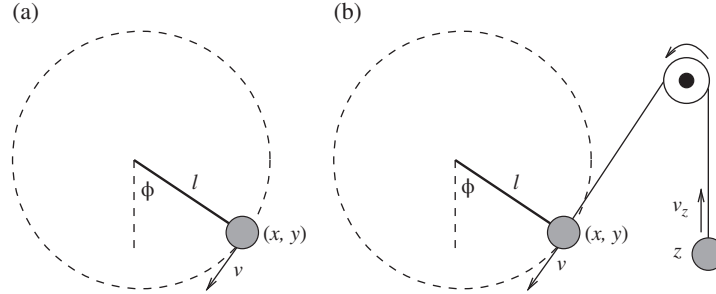
Figure 1.3 (a) Pendulum as a constrained problem; (b) coupled pendulum–wheel–mass system, where transformation into polar coordinates does not simplify the calculation.

1.2 Dynamics of rectilinear degrees of freedom

Labeling the coordinates with different letters such as x_1, y_1, z_1, ... will soon become inconvenient, so let us rename them as follows: $x_1 = r_1^{(1)}$, $y_1 = r_1^{(2)}$, $z_1 = r_1^{(3)}$, $x_2 = r_2^{(1)}$, $y_2 = r_2^{(2)}$, $z_2 = r_2^{(3)}$, ..., where the lower index represents the particle and the upper index in parentheses represents the dimension. The corresponding velocities can then be obtained as time derivatives:

$$v_i^{(j)} = \frac{\mathrm{d}}{\mathrm{d}t} r_i^{(j)} = \dot{r}_i^{(j)}.$$

If all the velocities vanish, we say that the system is *static*; if the velocities (which may be non-zero) do not change, we say that the system is *stationary*. The accelerations are the time derivatives of the velocities, or the second derivatives of the positions with respect to time: $a_i^{(j)} = \dot{v}_i^{(j)} = \ddot{r}_i^{(j)}$. If the acceleration is constant, we also refer to it as 'uniform'; in this case the velocity changes at a constant rate. For a particle i with mass m_i, Newton's equation of motion[1] expresses the relationship between the force $F_i^{(j)}$ applied to the particle and the acceleration $a_i^{(j)}$ in coordinate j as

$$\begin{aligned} F_i^{(j)} &= m_i \ddot{x}_i^{(j)} \\ &= m_i a_i^{(j)}. \end{aligned} \tag{1.6}$$

Numerical analysis prefers to deal with first-order equations, so often it is necessary to rewrite the second-order equation (1.6) as a first-order system by defining the velocity as an auxiliary variable:

$$F_i^{(j)} = m_i \dot{v}_i^{(j)}, \tag{1.7}$$

$$v_i^{(j)} = \dot{x}_i^{(j)}. \tag{1.8}$$

---

[1] This second-order differential equation formulation is actually due to Euler. Newton wrote his second law of motion as a first-order differential equation $F = \dot{p}$, where p is the momentum, but mathematically this is not equivalent to Euler's formulation.

Thus, instead of $3n$ second-order equations for n particles in three dimensions, we end up with $6n$ first-order equations. So, for a mechanical problem, one can choose whether to describe the system using first- or second-order differential equations. Consequently, physicists tend to call any equation with a first- or second-order time derivative on one side an 'equation of motion'. For example, the quantum-mechanical wave equations are called 'equations of motion of the probability' due to their relation with probability densities [1], and the time-dependent heat equation is sometimes called the 'equation of motion of heat' [2].

1.3 Dynamics of angular degrees of freedom

1.3.1 Rotation in two dimensions

In two dimensions, we have three degrees of freedom: two for translation and one for rotation. Rotation of a vector $r = (x, y)^{\mathrm{T}}$ by an angle ϕ in the xy-plane is represented by the rotation matrix for counterclockwise rotations,

$$A^{\phi} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \tag{1.9}$$

so a rotated vector r becomes

$$r' = A^{\phi}r = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix}. \tag{1.10}$$

The inverse transformation of a rotation by ϕ is represented by the transpose of the original rotation matrix. That the inverse is equal to the transpose characterizes an orthogonal matrix, a matrix whose columns are orthogonal to each other (i.e. have scalar product zero). The determinant of an orthonormal matrix is 1, so the length of a vector r which is rotated using a matrix of the form (1.9) does not change, and if two different vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are rotated into
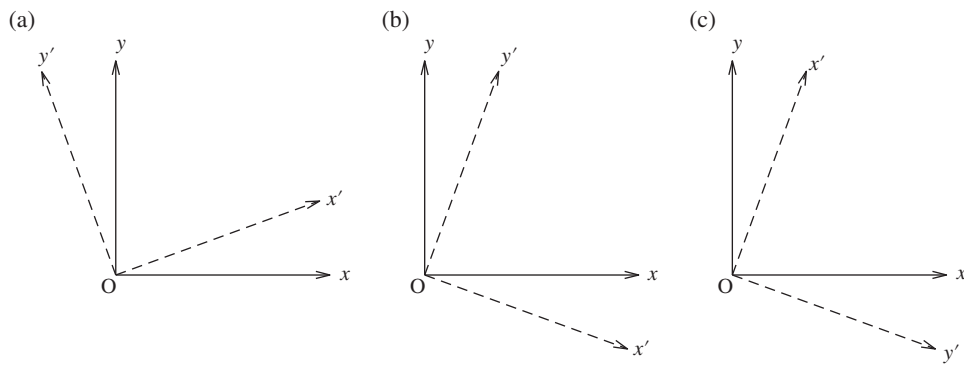
Figure 1.4 Original unit vectors x, y and transformed unit vectors x', y' for: (a) counterclockwise rotation by $20°$; (b) clockwise rotation by $20°$; (c) a combination of reflection in the $y = x$ line, such that the order of x and y is interchanged, and counterclockwise rotation.

$\tilde{\mathbf{r}}_1$ and $\tilde{\mathbf{r}}_2$, the angle between them will not change; see Figure 1.4(a). Matrices for clockwise rotation are of the form

$$B_\phi = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{1.11}$$

and also have determinant 1. Reflections, such as the one represented by the matrix

$$R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.12}$$

have determinant -1, as do compositions of reflections and rotations. In these cases, the angles between the transformed vectors $\tilde{\mathbf{r}}_1$, $\tilde{\mathbf{r}}_2$ and the original vectors $\mathbf{r}_1$, $\mathbf{r}_2$, as well as their relative orientations in the coordinate system, will change; see Figure 1.4(c).

1.3.2 Moment of inertia

With ϕ in (1.9) as the degree of freedom, the associated velocity will be the angular velocity $\omega = \dot{\phi}$. Let us next derive the inertia associated with the angular velocity ω via the kinetic energy, by way of a 'thought experiment' (Figure 1.5) where again we introduce constraints. Suppose that a point mass is initially moving at constant velocity in a straight trajectory, as shown in Figure 1.5(a); then its kinetic energy will be $E_{\text{kin}} = \frac{1}{2}mv^2$. Now let the mass be captured by a constraint which is connected to a point at fixed distance r from the trajectory, as in Figure 1.5(b); assuming no energy dissipation, the point mass will then rotate with the same kinetic energy as before; see Figure 1.5(c). With $v = r\omega$, from the kinetic energy expression for the rectilinear degree of freedom we have

$$\frac{1}{2}mv^2 = \frac{1}{2}mr^2\omega^2. \tag{1.13}$$
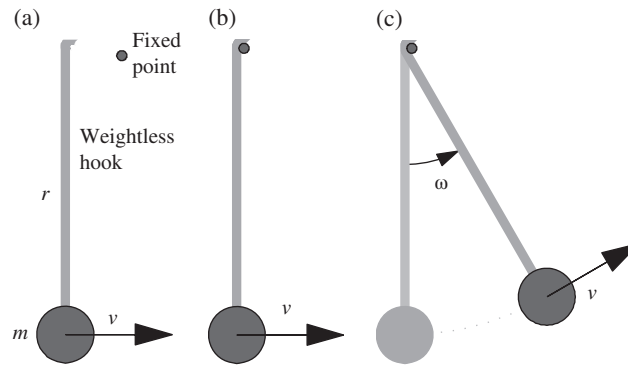
Figure 1.5 Thought experiment: (a) a mass m at the end of a weightless rigid hook is initially in linear motion; (b) the hook is caught by a fixed point, without any change in the kinetic energy; (c) the mass m starts to circle the fixed point with angular motion.

So the moment of inertia for the rotation is $J = mr^2$ for a point mass. In analogy to the linear momentum $p = mv$, which can be used to write the kinetic energy as $E_{\text{kin,lin}} = p^2/(2m)$, for angular motion we have the angular momentum $L = J\omega$ and the kinetic energy

$$E_{\text{kin,rot,2D}} = L^2/(2J).$$

A frequent source of error in calculations is use of the frequency f instead of the angular velocity

$$\omega = 2\pi f,$$

where f is computed directly from the number of turns N_{turn} around the center of rotation during time T as

$$f = \frac{N_{\text{turn}}}{T}.$$

Using f when ω should be used introduces an error of factor $(2\pi)^2 \approx 40$ in calculations of the kinetic energy.

We can compute the kinetic energy of a mass point for either the rectilinear or the angular degree of freedom. For extended masses, we have to use the moment of inertia with respect to the center of mass (König's theorem). For a body made up of n equal point masses m at positions r_i, the moment of inertia for a rotation around the center of mass $\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$ is

$$J_{\text{e}} = m \sum_{i=1}^{n}(r_i - \bar{r})^2$$

(where the subscript 'e' stands for 'equal'). For unequal point masses m_i, the arithmetic mean for the center of mass is replaced by the weighted average

$$\bar{r} = \frac{\sum_{i=1}^{n} r_i m_i}{\sum_{i=1}^{n} m_i},$$

and the moment of inertia for a body composed of these masses is

$$J_{\text{ue}} = \sum_{i=1}^{n} m_i(r_i - \bar{r})^2 \tag{1.14}$$

(where the subscript 'ue' stands for 'unequal'). For a body with continuous mass distribution, the summation in (1.14) should be replaced by an integral

$$J = \int r^2 \, dm \tag{1.15}$$

for the distance r of the infinitesimal mass element dm which depends on r.

In two-dimensional DEM simulations for homogeneous bodies, we may use a two-dimensional density σ (with units kg/m^2) obtained from three-dimensional densities ρ (with units kg/m^3) upon multiplying by the depth of the system. With meter-long rods as particles,

the numerical value of the three-dimensional density can be used. The calculation then simplifies to

$$J = \sigma \int r^2 \, \mathrm{d}A(r). \tag{1.16}$$

As can be seen from this formula, the moment of inertia depends on the distance r from the position of the rotation axis, so the same body will have a different moment of inertia for different positions of the rotation axis; for example, a stick held at one end has a larger moment of inertia than if it was held in the middle. In general, moments of inertia are tabulated for the center of mass.

For a body with mass m and a rotation axis through an arbitrary point $\mathbf{r}_\mathrm{r}$ which is not the center of mass $\mathbf{r}_\mathrm{c}$, we can compute its moment of inertia via the parallel axis theorem (Steiner's or the Huygens–Steiner theorem),

$$J_\mathrm{r} = J_\mathrm{c} + m \|\mathbf{r}_\mathrm{c} - \mathbf{r}_\mathrm{r}\|^2, \tag{1.17}$$

where the norm (length) $\|\cdot\|$ of a vector $\mathbf{r} = (r_1, r_2)$ is given by

$$\|\mathbf{r}\| = \sqrt{r_1^2 + r_2^2} = r. \tag{1.18}$$

Thus the moment of inertia is equal to the sum of the moments of inertia, plus the moment of inertia of the center of mass with respect to the rotation axis. For particles of (convex) polygonal shape in two dimensions, a decomposition into non-overlapping triangles with one corner at the center of mass of the polygon is the easiest way to compute the moment of inertia. With respect to the center of mass, the moment of inertia of a triangular mass with base b and height h is given by
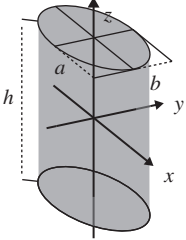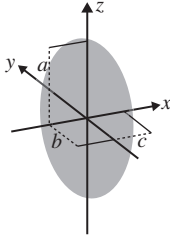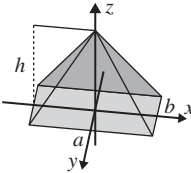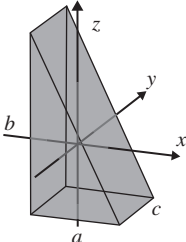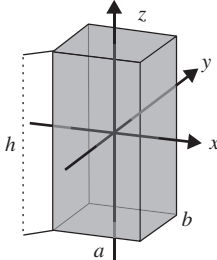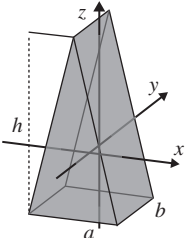
$$J^\triangle = \sigma \frac{bh^3}{36},$$

where σ is the two-dimensional density. The height h can be computed as the length of the vector rejection of the position vector of the corner opposite the base b. The moment of inertia with respect to the corner point can then be obtained by shifting $J^\triangle$ according to Equation (1.17). For conventional homogeneous two-dimensional bodies, the moments of inertia for a rotation around the center of mass in the z-direction can be obtained from Table 1.1 for three-dimensional bodies by setting the height in the third dimension to zero.

1.3.3 *From two to three dimensions*

While rotations in two dimensions with only the z-axis as rotation axis behave very much like rectilinear degrees of freedom under translations, in three dimensions the dynamics is different. When a rigid object described by points $\mathbf{r}_i$ undergoes a translation $\mathcal{T}_{\delta\mathbf{r}}$, the positions of all points change by the same amount $\delta\mathbf{r} = \delta x + \delta y + \delta z$, so that the new points are $\tilde{\mathbf{r}}_i = \mathbf{r}_i + \delta\mathbf{r}$. We can decompose the translation into axis-parallel components $\mathcal{T}_{\delta x}$, $\mathcal{T}_{\delta y}$ and $\mathcal{T}_{\delta z}$, which all commute:

$$\mathcal{T}_{\delta\mathbf{r}} = \mathcal{T}_{\delta x} + \mathcal{T}_{\delta y} + \mathcal{T}_{\delta z} = \mathcal{T}_{\delta x} + \mathcal{T}_{\delta z} + \mathcal{T}_{\delta y} = \cdots,$$

**Table 1.1** Moments of inertia [5] for various solid bodies with homogeneous density distribution and rotation axis through the center of mass; a, b and c denote the half-axes for the cylinder and ellipsoid but full side-lengths for the other figures.

Elliptical cylinder	Ellipsoid	Rectangular pyramid
$J_x = \dfrac{m}{12}(3b^2 + h^2)$	$J_x = \dfrac{m}{5}(a^2 + b^2)$	$J_x = \dfrac{m}{20}\left(b^2 + \dfrac{3h^2}{4}\right)$
$J_y = \dfrac{m}{12}(3a^2 + h^2)$	$J_y = \dfrac{m}{5}(a^2 + c^2)$	$J_y = \dfrac{m}{20}\left(a^2 + \dfrac{3h^2}{4}\right)$
$J_z = \dfrac{m}{4}(a^2 + b^2)$	$J_z = \dfrac{m}{5}(b^2 + c^2)$	$J_z = \dfrac{m}{20}(a^2 + b^2)$

Wedge with right angle	Solid cuboid	Isoceles (symmetric) wedge
$J_x = \dfrac{m}{36}(2h^2 + 3b^2)$	$J_x = \dfrac{m}{12}(b^2 + h^2)$	$J_x = \dfrac{m}{36}(2h^2 + 3a^2)$
$J_y = \dfrac{m}{18}(a^2 + h^2)$	$J_y = \dfrac{m}{12}(a^2 + h^2)$	$J_y = \dfrac{m}{72}(4h^2 + 3b^2)$
$J_z = \dfrac{m}{36}(2a^2 + 3b^2)$	$J_z = \dfrac{m}{12}(a^2 + b^2)$	$J_z = \dfrac{m}{24}(2a^2 + b^2)$

i.e. the order of componentwise translations can be interchanged without changing the result. Just writing the translation in Cartesian coordinates shows that it has the nature of vector addition. Rotations, however, are different. Let us write a rotation around the x-axis by angle ϕ as $R_x(\phi)$ and a rotation around the y-axis by angle θ as $R_y(\theta)$; do $R_x(\phi)$ and $R_y(\theta)$ commute? What happens if we try out 90° rotations of a book? In Figure 1.6 we see that we get different final positions depending on the order in which we perform the rotations. So one has to be much more careful with rotations than with translations.
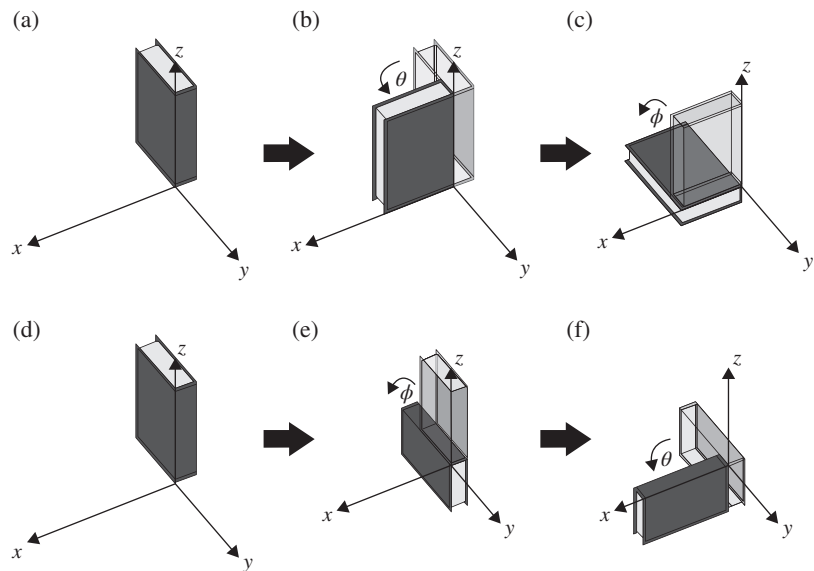
Figure 1.6 Rotations of a book by 90° around two axes: in the sequence from (a) to (c), the book is rotated first around the z-axis and then around the x-axis; in the sequence from (d) to (f), the book is rotated first around the x-axis and then around the z-axis.
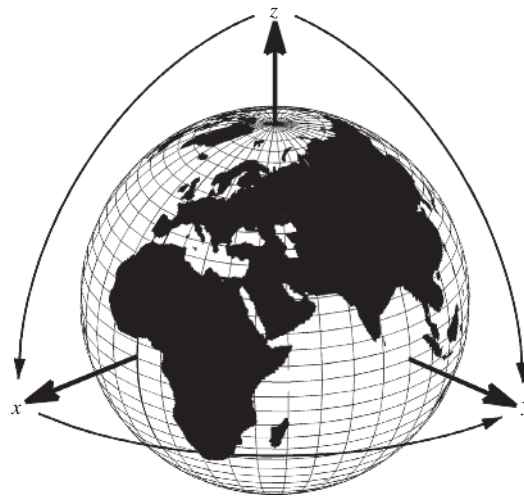
Figure 1.7 Rotation around two of three orthogonal axes (x and y) on a globe gives a rotation around the third axis (z).

In Figure 1.7, when we rotate the positive x-axis by 90° around the y-axis from the equator to the north pole, and then down again by 90° around the x-axis, we get the same result as if we had simply rotated it by 90° around the z-axis. Thus, rotations as degrees of freedom in three dimensions behave very differently from translations: not only do they not commute
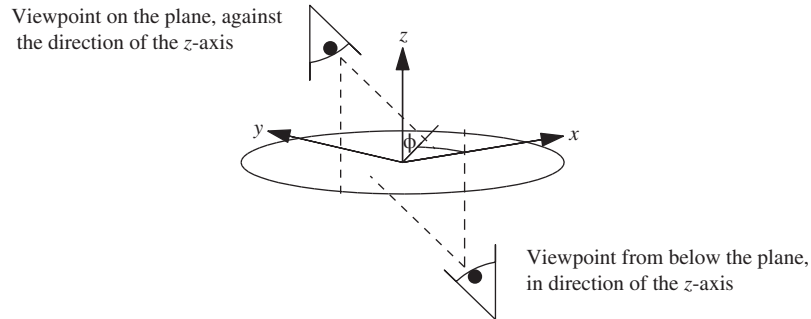
Figure 1.8   Rotations in the xy-plane: viewed from above the origin against the direction of the z-axis, with the angle ϕ defined counterclockwise, a perspective which is employed in the two-dimensional formulation of the rotation matrix; and viewed from below the origin in the direction of the z-axis, a perspective employed in the three-dimensional formulation of the rotation matrix.

(the result is different if the order of the operations is changed), but two rotations around two orthogonal axes may have an identical effect to a single rotation around the third axis.

1.3.4   Rotation matrix in three dimensions

In both two and three dimensions, rotations are conventionally measured counterclockwise, but the meaning of 'counterclockwise' depends on the viewpoint one takes. The conventional view in two-dimensional geometry is from above the xy-plane, against the direction of the z-axis towards the origin; however, in three dimensions, this is the same as a clockwise rotation viewed from the origin in the direction of the positive z-axis; see Figure 1.8. This means that a rotation matrix written as

$$A^\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{1.19}$$

in two dimensions is equivalent to the same rotation in the xy-plane around the z-axis in three dimensions, but in the opposite direction; so it is written as [3, 4]

$$\mathbf{A}^\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.20}$$

Although it looks as if the sign or direction of the angle ϕ is reversed, in fact it is the viewpoint that is reversed. Apart from this difference in sign, the three-dimensional rotation matrix around the z-axis is obtained by simply augmenting the two-dimensional rotation matrix by putting a 1 in the lower right diagonal and zeros elsewhere, because this rotation leaves the z-components unchanged.

While for translations the order of operations does not matter, for rotations in three dimensions we have to specify the order of the elementary operations. Using the classical convention of Euler angles, the parametrization of the three degrees of freedom is as follows: first, rotation

by ϕ around the z-axis; second, rotation by θ around the x-axis; third, rotation by ψ around the new z-axis. The rotation matrix around the x-axis by θ has 1 as its upper left entry:

$$\mathbf{A}^\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

The final rotation by angle ψ again looks like $\mathbf{A}^\phi$:

$$\mathbf{A}^\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

but it should not be forgotten that this rotation is around the *new* z-axis, i.e. ψ is measured differently than ϕ. Together, the concatenated rotations give the rotation matrix

$$\mathbf{A} = \mathbf{A}^\psi \mathbf{A}^\theta \mathbf{A}^\phi$$
$$= \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & -\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi & \sin\theta\cos\psi \\ \sin\phi\sin\theta & -\cos\phi\sin\theta & \cos\theta \end{pmatrix}.$$
$$(1.21)$$

Three-dimensional vectors are transformed by $\mathbf{A}$ into the rotated coordinate system via

$$r' = \mathbf{A}r, \tag{1.22}$$

and the reverse transformation is given by the transpose:

$$r = \mathbf{A}^{-1}r' = \mathbf{A}^{\mathrm{T}}r'. \tag{1.23}$$

1.3.5 Three-dimensional moments of inertia

Let us derive the three-dimensional moment of inertia via the kinetic energy [6], using reasoning analogous to that in § 1.3.2 for the two-dimensional case, where the kinetic energy for a particle of mass m was

$$E^{\mathrm{kin,rot,2D}} = \frac{1}{2}m(\omega r)^2.$$

In three dimensions, we consider the moment of inertia for the rotation around $\boldsymbol{\omega}$ of a vector $\mathbf{r}$ which is orthogonal to $\boldsymbol{\omega}$, so the cross product $\times$ is needed:

$$E^{\mathrm{kin,rot,3D}} = \frac{1}{2}m(\boldsymbol{\omega} \times \mathbf{r})^2.$$

For a continuous mass distribution, with infinitesimal mass element $\mathrm{d}m$, this corresponds to

$$E^{\mathrm{kin,rot,3D}} = \frac{1}{2}\int (\boldsymbol{\omega} \times \mathbf{r})^2 \, \mathrm{d}m. \tag{1.24}$$

We shall use the following identities from vector analysis, which are valid for any vectors $\boldsymbol{\omega}$ and $\mathbf{r}$ (note that here and throughout the text we do not distinguish between row and column vectors, as this makes no sense for the cross product):

$$(\boldsymbol{\omega} \times \mathbf{r})^2 = \boldsymbol{\omega}\,(\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})),$$

$$\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) = \left(\mathbf{r}^2 \mathbb{1} - \mathbf{rr}\right)\boldsymbol{\omega},$$

where $\mathbb{1}$ denotes the three-dimensional identity matrix. Using these identities, we can rewrite (1.24) and extract $\boldsymbol{\omega}$ from under the integral to get

$$\frac{1}{2}\int (\boldsymbol{\omega} \times \mathbf{r})^2 \, \mathrm{d}m = \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\left(\int (\mathbf{r}^2\mathbb{1} - \mathbf{rr})\,\mathrm{d}m\right)\boldsymbol{\omega}. \tag{1.25}$$

Thus, the integral turns out to be the tensor of the moment of inertia

$$\mathbf{J} = \int (\mathbf{r}^2\mathbb{1} - \mathbf{rr})\,\mathrm{d}m.$$

Writing the tensor in component form, with $\mathbf{r} = (x, y, z)$, yields

$$\mathbf{J} = \begin{pmatrix} \int_m (y^2 + z^2)\,\mathrm{d}m & -\int_m xy\,\mathrm{d}m & -\int_m xz\,\mathrm{d}m \\ -\int_m xy\,\mathrm{d}m & \int_m (x^2 + z^2)\,\mathrm{d}m & -\int_m xz\,\mathrm{d}m \\ -\int_m xz\,\mathrm{d}m & -\int_m xz\,\mathrm{d}m & \int_m (x^2 + y^2)\,\mathrm{d}m \end{pmatrix}$$

$$= \begin{pmatrix} J_{xx} & -J_{xy} & -J_{xz} \\ -J_{xy} & J_{yy} & -J_{yz} \\ -J_{xz} & -J_{yz} & J_{zz} \end{pmatrix}.$$

The axes of the coordinate system in which the tensor $\mathbf{J}$ is diagonal are called the principal axes of the moment of inertia, and the diagonal elements J_1, J_2 and J_3 are the eigenvalues of $\mathbf{J}$. The angular velocity in three dimensions is $\boldsymbol{\omega}$, whose z-component ω_z we used (without the subscript z) in the previous section as the two-dimensional angular velocity. The angular momentum is

$$\mathbf{L} = \mathbf{J}\boldsymbol{\omega}.$$

The kinetic energy for the rectilinear degrees of freedom of a particle in three dimensions is

$$E_{\mathrm{rli}} = \frac{1}{2}m|\mathbf{v}|^2, \tag{1.26}$$

where $|\mathbf{v}|^2$ is the scalar product of the velocity with itself. The associated kinetic energy for three-dimensional angular motion is a quadratic form of the tensor $\mathbf{J}$ and the vectorial angular velocity $\boldsymbol{\omega}$:

$$E_{\mathrm{rot}} = \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\mathbf{J}\boldsymbol{\omega}. \tag{1.27}$$

The parallel axis theorem (also known as Steiner's theorem or the Huygens–Steiner theorem) says that if J_{ij}^{cm} is the inertia tensor with respect to the center of mass, then the inertia tensor

relative to a point separated by $\vec{a} = (a_1, a_2, a_3)$ from the center of mass along the same (principal) axes is

$$J_{ij} = J_{ij}^{\text{cm}} + m \left(|\vec{a}|^2 \delta_{ij} - a_i a_j \right), \tag{1.28}$$

where δ_{ij} is the Kronecker delta,

$$\delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Written explicitly in component form, this is

$$\mathbf{J} = \mathbf{J}^{\text{cm}} + m \begin{pmatrix} a_2^2 + a_3^2 & -a_1 a_2 & -a_1 a_3 \\ -a_1 a_2 & a_1^2 + a_3^2 & -a_2 a_3 \\ -a_1 a_3 & -a_2 a_3 & a_1^2 + a_2^2 \end{pmatrix}. \tag{1.29}$$

If the axis of rotation is not along the principal axis, the tensor transforms via the rotation matrix $\mathbf{A}$ to

$$\mathbf{J} = \mathbf{A} \mathbf{J}^{\text{cm}} \mathbf{A}^{\text{T}}. \tag{1.30}$$

Useful three-dimensional moments of inertia with rotation axis through the center of mass are given in Table 1.1. For polyhedral particles, it is best to use a successive decomposition into tetrahedra; see § 8.2.3. The corresponding two-dimensional moments of inertia are obtained by manipulating the other axis dimensions. For a thin stick of length h, the moment of inertia for rotation around the x-axis can be obtained from the formula for the cylinder by setting r to 0, so that one obtains $(m/12)h^2$; the moment of inertia of a sphere can be obtained from the formula for the ellipsoid by setting all semi-axes equal to the radius, and so on. For rotation axes not through the center of mass, the moment of inertia can be computed via the parallel axis theorems in two and three dimensions, namely Equations (1.17) and (1.28) or (1.29).

The joint moments of inertia $\mathbf{J}^{\text{j}}$ of a shape s1 that is hollowed out by another shape s2 can be obtained by subtracting the individual moments of inertia:

$$\mathbf{J}^{\text{j}} = \mathbf{J}^{\text{s1}} - \mathbf{J}^{\text{s2}}, \tag{1.31}$$

which is simple if one only wants to obtain numerical values. However, when one wants to obtain scaling by the mass m of the new, hollow body, the derivations become more cumbersome. For a cylinder with outer radius r_0, we obtain its moment of inertia from the formula for the elliptical cylinder in Table 1.1 as $J_z = \frac{1}{2} m_0 r_0^2$. When the cylinder is hollowed out concentrically in the middle by a cylindrical shape so that the radius of the inner hollow is r_{i}, the mass becomes

$$m^{\text{hc}} = m_0 \frac{r_0^2 - r_{\text{i}}^2}{r_0^2} = m_0 - m_{\text{i}},$$

where m_{i} is the mass which has been removed from the original cylinder. So, for the moment of inertia along the symmetry axis, one obtains

$$J^{\mathrm{hc}} = \frac{1}{2}m_0 r_0^2 - \frac{1}{2}m_i r_i^2$$

$$= \frac{1}{2}m_0 r_0^2 - \frac{1}{2}\left(m_0 - m_0\frac{r_0^2 - r_i^2}{r_0^2}\right)r_i^2$$

$$= \frac{1}{2}m_0\frac{r_0^4}{r_0^2} - \frac{1}{2}m_0\frac{r_i^4}{r_0^2}$$

$$= \frac{m^{\mathrm{hc}}}{2}\frac{r_0^2}{r_0^2 - r_i^2}\left(\frac{r_0^4}{r_0^2} - \frac{r_i^4}{r_0^2}\right)$$

$$= \frac{m^{\mathrm{hc}}}{2}\left(r_0^2 + r_i^2\right).$$

Notice that the sum, not the difference, of the squares of the radii appears in the final formula.

1.3.6 Space-fixed and body-fixed coordinate systems and equations of motion

To describe the dynamics (equations of motion) of a three-dimensional system, we will use two particular coordinate systems in our computer simulations: space-fixed coordinate systems, where vectors $\mathbf{e}^s$ are expanded in the unit vectors (e_x, e_y, e_z) of the Cartesian 'laboratory frame'; and body-fixed coordinate systems, where basis vectors $\mathbf{e}^b$ aligned with the body are expanded in unit vectors (e_1, e_2, e_3) chosen so that the origin is at the center of mass of the body and their orientations are such that the tensor of inertia is diagonal; see Figure 1.9.

Torques are usually computed in space-fixed coordinates (i.e. an unrotated coordinate system without inertia forces); we denote these by $\boldsymbol{\tau}^s$, and they equal the rate of change with respect to time of the moment of inertia:

$$\dot{\mathbf{J}}^s = \boldsymbol{\tau}^s.$$
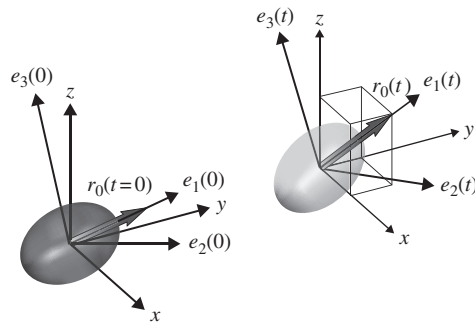
Figure 1.9 Cartesian laboratory frame with axes (x, y, z) and body-fixed coordinate system (e_1, e_2, e_3) at time 0 and at time t; also shown is the space-fixed orientation of the body, r_0, at time 0 and at time t when the body is at a different position.
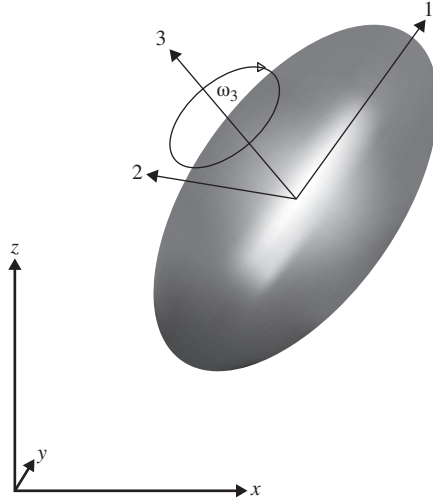
Figure 1.10 Constant angular velocity ω_3 around the e_3-axis of a body-fixed coordinate system with axes (e_1, e_2, e_3) and a Cartesian laboratory frame with axes (x, y, z). The angular velocity $(\omega_1, \omega_2, \omega_3)$ is finite but does not correspond to the time derivative of an angular coordinate in the body-fixed system.

To obtain the corresponding equations in body-fixed (embedded) coordinates, we have to explicitly compute the transformations to factor in the inertia forces due to rotation. We can then obtain the space-fixed angular velocities $\boldsymbol{\omega}^s = (\omega_x^s, \omega_y^s, \omega_z^s)$ from the rotation matrix $\mathbf{A}$ via Equation (1.23):

$$\boldsymbol{\omega}^s = \mathbf{A}^{-1}\boldsymbol{\omega}^b = \mathbf{A}^T\boldsymbol{\omega}^b.$$

Just because the coordinate system rotates with the body does not mean that the angular velocity in the body-fixed coordinate system is zero. Even if there is conservation of angular momentum, so that $\mathbf{L}^b = \mathbf{J}^b\boldsymbol{\omega}^b = $ constant, it merely implies that the $\boldsymbol{\omega}^b$ would be constant, too; see Figure 1.10. Nevertheless, this angular velocity does not correspond to the change of an angle in the body-fixed coordinate system. For the particular body-fixed coordinate system where the tensor of inertia is diagonal, i.e.

$$\mathbf{J}^b = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_2 \end{pmatrix},$$

we obtain the rate of change of the moment of inertia via the product rule:

$$\begin{aligned}
\dot{\mathbf{j}}^b &= \frac{d}{dt}\left(J_1\omega_1 e_1 + J_2\omega_2 e_2 + J_3\omega_3 e_3\right) \\
&= \left(J_1\dot{\omega}_1 e_1 + J_2\dot{\omega}_2 e_2 + J_3\dot{\omega}_3 e_3\right) + \left(J_1\omega_1\dot{e}_1 + J_2\omega_2\dot{e}_2 + J_3\omega_3\dot{e}_3\right).
\end{aligned} \tag{1.32}$$
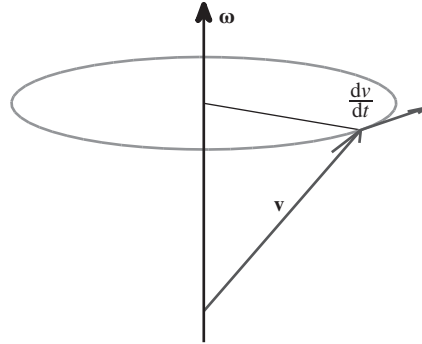
Figure 1.11   For the rotation of a vector **v** from the origin of the body-fixed coordinate system with angular velocity $\boldsymbol{\omega}$, the time derivative $d\mathbf{v}/dt$ is orthogonal to both **v** and $\boldsymbol{\omega}$.

For the vectors that are fixed in the body-fixed coordinate system, the time derivative is

$$\dot{\mathbf{e}}^{b} = \boldsymbol{\omega}^{b} \times \mathbf{e}^{b}, \tag{1.33}$$

i.e. the changes of vectors rotating with the body are orthogonal to the angular velocity $\boldsymbol{\omega}$ and the vector **e** itself; see Figure 1.11. The torques $\boldsymbol{\tau}^{b}$ in the body-fixed coordinate system can be obtained from the rotation matrix (1.21) as

$$\boldsymbol{\tau}^{b} = \mathbf{A}\boldsymbol{\tau}^{s}. \tag{1.34}$$

From Equations (1.32) and (1.33) one then obtains the 'Euler equations of motion'

$$\tau_{1}^{b} = J_{1}\dot{\omega}_{1}^{b} - (J_{2} - J_{3})\,\omega_{2}^{b}\omega_{3}^{b}, \tag{1.35}$$

$$\tau_{2}^{b} = J_{2}\dot{\omega}_{2}^{b} - (J_{3} - J_{1})\,\omega_{1}^{b}\omega_{3}^{b}, \tag{1.36}$$

$$\tau_{3}^{b} = J_{3}\dot{\omega}_{3}^{b} - (J_{1} - J_{2})\,\omega_{1}^{b}\omega_{2}^{b}, \tag{1.37}$$

for the angular velocities around the respective body-fixed axes $e_{1}^{b}, e_{2}^{b}, e_{3}^{b}$ through the center of mass. The corresponding Newton equations for the velocities v_{i} of the rectilinear degrees of freedom,

$$F_{i} = m\dot{v}_{i}, \tag{1.38}$$

are linear in the v_{i}, whereas the Euler equations (1.35)–(1.37) for the rotations are nonlinear in the ω_{i} (except for bodies with equal moments of inertia along the principal axes, like spheres and symmetric polyhedra), since successive rotations around two axes are equivalent to a rotation around the third axis, as illustrated in Figure 1.7. For rotations around a single axis only, two components in (1.35)–(1.37), ω_{1} and ω_{2} say, will vanish, and the remaining equation becomes linear in the third component, ω_{3}, and is equivalent to the Newton equation (1.38). When the nonlinear terms in the Euler equations (1.35)–(1.37) are neglected, the particle trajectories will be different, i.e. wrong. An example for a polygon rolling down a slope is given in Figure 1.12. Newton's equation of motion (1.38) and the Euler equations of motion (1.35)–(1.37) are sometimes collectively called the Newton–Euler equations of motion, to emphasize the necessity of including the angular degrees of freedom appropriately.
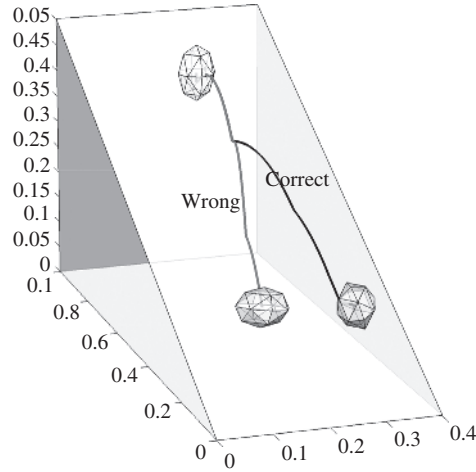
Figure 1.12 Correct trajectory (black, with Euler equations of motion) and incorrect trajectory (gray, neglecting nonlinear terms in the Euler equations) computed over the time interval $t \in [0, 0.6]$ s for a particle with initial angular velocity $\boldsymbol{\omega} = 10\pi(1, 1, 1)\,\mathrm{s}^{-1}$ and initial velocity $v = (0, 0, 0.025)\,\mathrm{m\,s}^{-1}$.

1.3.7 Problems with Euler angles

The inverse of **A** is obtained by reversing the order of multiplication of $\mathbf{A}^\psi$, $\mathbf{A}^\theta$ and $\mathbf{A}^\phi$ and using the opposite angle:

$$\mathbf{A}^{-1} = \mathbf{A}^{-\phi}\mathbf{A}^{-\theta}\mathbf{A}^{-\psi}.$$

This turns out to be exactly equal to the transpose of **A**, so that $\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}$. As with two-dimensional rotations, **A** is an orthogonal matrix, a property which is inherited from the matrices $\mathbf{A}^\psi$, $\mathbf{A}^\theta$ and $\mathbf{A}^\phi$. That all is not well with Euler angles can be seen when the angles ψ, θ, ϕ are again computed from a given rotation matrix

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix}.$$

One obtains the angles ψ, θ, ϕ from the $A_{i,j}$ through the following relations:

$$\cos\theta = A_{3,3}, \qquad \sin\theta = \pm\sqrt{1 - \cos^2\theta},$$

$$\cos\psi = -\frac{A_{3,2}}{\sin\theta}, \qquad \sin\psi = -\frac{A_{3,1}}{\sin\theta},$$

$$\cos\theta = \frac{A_{2,3}}{\sin\theta}, \qquad \sin\phi = -\frac{A_{1,3}}{\sin\theta}.$$

These equations show that a position close to $\theta = \pi/2$ will lead to a divergence in the computation of the Euler angles. Worse is to come when we discuss the equations of motion for ϕ, θ and ψ. For the time derivatives $\dot{\phi}, \dot{\theta}, \dot{\psi}$ of the angular coordinates, we obtain (by writing

$\boldsymbol{\omega}$ in terms of ϕ, θ, ψ and their time derivatives and then solving the resulting three equations for $\dot{\phi}$, $\dot{\theta}$, $\dot{\psi}$; see Greenwood ([7], p. 144)

$$\dot{\phi} = -\omega_x^s \frac{\sin\phi\cos\theta}{\sin\theta} + \omega_y^s \frac{\cos\phi\cos\theta}{\sin\theta} + \omega_z^s, \tag{1.39}$$

$$\dot{\theta} = \omega_x^s \cos\phi + \omega_y^s \sin\phi, \tag{1.40}$$

$$\dot{\psi} = \omega_x^s \frac{\sin\phi}{\sin\theta} - \omega_y^s \frac{\cos\phi}{\sin\theta}. \tag{1.41}$$

For $\theta = 0$, the equations of motion for the Euler angles become singular, too! This is definitely bad news when we run simulations of many particles which are not constrained whatsoever: the likelihood that an equation for some particle diverges because it comes close to $\theta = 0$ increases with the number of particles and the simulation time. (The widespread use of Euler angles and their equations of motion in theoretical mechanics textbooks is due to the fact that for many single-particle problems, such as a top constrained at its tip, the singularities can be eliminated by a suitable choice of variables.) Allen and Tildesley [4] propose a possible workaround that uses two coordinate systems, performing the time integration in the one where the particle is away from $\theta = 0$, but we prefer to use quaternions (see § 1.3.9), for which singularities in the time integration do not occur at all.

1.3.8 Rotations represented using complex numbers

In this subsection we briefly recall the basic properties of complex numbers, as an prelude to quaternions. With the imaginary unit $i = \sqrt{-1}$, we can define complex numbers $z \in \mathbb{C}$ in the form $z = x + iy$ where $x, y \in \mathbb{R}$. Two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ behave exactly as vectors in $\mathbb{R}^2$ when it comes to equality, addition and subtraction—complex numbers 'are' two-dimensional vectors. What differs is the product: instead of the inner product for two-dimensional vectors, which results in a smaller dimensionality than the original $\mathbb{R}^2$ vectors, complex multiplication gives

$$z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

With the complex conjugate $\bar{z} = x - iy$ of $z = x + iy$, one can write the absolute value of a complex number (i.e. the length of the corresponding two-dimensional vector) as

$$|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}. \tag{1.42}$$

This also allows us to reduce division of two complex numbers to real divisions as follows:

$$\frac{z_1}{z_2} = \frac{z_1}{z_2}\frac{\bar{z}_2}{\bar{z}_2} = \frac{x_1 + iy_1}{x_2 + iy_2}\frac{x_2 - iy_2}{x_2 - iy_2}$$
$$= \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + i\frac{-x_1 y_2 + x_2 y_1}{x_2^2 + y_2^2}.$$

As an alternative to Equations (1.9) and (1.10), we can rewrite the two-dimensional vector r as complex number $z = x + iy$ and then, using the Euler formula for a complex exponential,

$$\exp(i\phi) = \cos\phi + i\sin\phi, \tag{1.43}$$
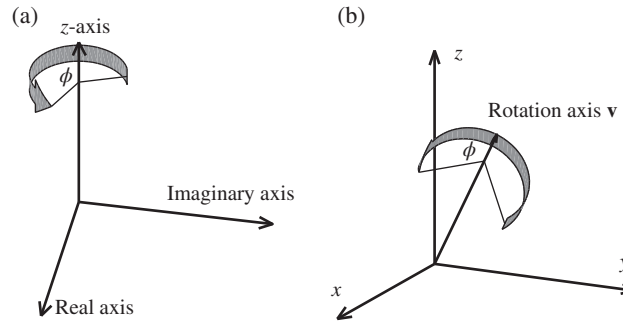
Figure 1.13 Rotation by angle ϕ: (a) in two dimensions around the z-axis, via the complex exponential $\exp(i\phi)$; (b) in three dimensions around an arbitrary rotation axis, via the unit quaternions (Euler parameters) $q = (\cos\phi, \hat{\mathbf{v}}\sin\phi)$ and $-q^* = (-\cos\phi, \hat{\mathbf{v}}\sin\phi)$. Angles are measured counterclockwise with respect to the axis of rotation (ϕ measured with respect to $\mathbf{v}$, $-\phi$ measured with respect to $-\mathbf{v}$). Diagram (b) and similar ones in the literature should be taken with a grain of salt: for Euler parameters, it is actually $\phi/2$ that is needed in the argument for a rotation by ϕ.

formulate the rotation in (1.10) as a complex multiplication; see Figure 1.13. Specifically, we see that

$$
\begin{aligned}
\exp(i\phi)z &= (\cos\phi + i\sin\phi)\,(x + iy) \\
&= (x\cos\phi - y\sin\phi) + i\,(y\cos\phi + x\sin\phi)
\end{aligned}
$$

is equivalent to the representation of vector rotation via matrix multiplication in Equation (1.10). Most notably, a multiplication by the imaginary unit i corresponds to a counterclockwise rotation by $90°$. The reason the complex exponential works so well as a replacement for the rotation matrix (1.9) is that its eigenvalues are $\exp(\pm i\phi)$.

1.3.9 Quaternions

There are many possible ways to represent three-dimensional rotation, but many are cumbersome[2] or the numerical implementations of their equations of motion are unstable; see the discussion following Equations (1.39)–(1.41). We shall use quaternions as the basis for representing three-dimensional rotations: quaternions are illuminating in their 'vector representation', and they can be implemented numerically to obtain very stable representations of the equations of motion for the angular degrees of freedom. As the complex exponential $\exp(i\phi)$ worked so nicely to describe rotations in two dimensions, we seek an extension of the concept to use for general rotations in three dimensions. While for the complex exponential only the z-axis was treated as the rotation axis, in three dimensions we will have to deal with arbitrary rotation axes, and more relations will be needed than for complex numbers. The

---

[2]For example, if one directly uses the rotation matrix, which has nine entries for the three independent rotational degrees of freedom, then six constraints must be introduced and integrated out; likewise for the corresponding time derivatives.

quaternion basis elements I, J and K are chosen to satisfy the relations

$$\mathsf{I} \cdot \mathsf{I} = -1, \quad \mathsf{J} \cdot \mathsf{J} = -1, \quad \mathsf{K} \cdot \mathsf{K} = -1, \tag{1.44}$$

just like the imaginary unit i. We will develop the meaning of quaternions a bit at a time. To begin with, we warn that the basis element I is not identical to the imaginary unit i, and the basis element J is a different variable from the tensor of inertia $\mathbf{J}$ and its components. Representations of I, J and K via real and complex matrices are explored in Exercise 1.2. We also need relations between I, J, K and their products:

$$\begin{aligned} \mathsf{I} \cdot \mathsf{J} &= \ \ \mathsf{K}, & \mathsf{J} \cdot \mathsf{K} &= \ \ \mathsf{I}, & \mathsf{K} \cdot \mathsf{I} &= \ \ \mathsf{J}, \\ \mathsf{J} \cdot \mathsf{I} &= -\mathsf{K}, & \mathsf{K} \cdot \mathsf{J} &= -\mathsf{I}, & \mathsf{I} \cdot \mathsf{K} &= -\mathsf{J}. \end{aligned} \tag{1.45}$$

In the upper line are the 'cyclic permutations' according to alphabetical order, i.e. multiplying together two of the basis elements gives the third; the lower line contains the 'anti-cyclic permutations', where multiplying together two of the basis elements gives minus the third. The upper line is reminiscent of Figure 1.7, where successive rotations around two of the orthogonal axes yielded rotation around the third axis. Comparing the two lines of (1.45) shows that reversal of the multiplication order is anti-commutative (i.e. the sign of the result is reversed if the order in the product is reversed); this is reminiscent of the cross product for vectors. Additionally, for quaternions there is a 'unit operation' $\mathbb{1}$, such that

$$\mathsf{I} \cdot \mathbb{1} = \mathbb{1} \cdot \mathsf{I} = \mathsf{I}, \quad \mathsf{J} \cdot \mathbb{1} = \mathbb{1} \cdot \mathsf{J} = \mathsf{J}, \quad \mathsf{K} \cdot \mathbb{1} = \mathbb{1} \cdot \mathsf{K} = \mathsf{K}. \tag{1.46}$$

Next, we use I, J, K and $\mathbb{1}$ to define a general quaternion $\mathbf{q}$ and its conjugate $\mathbf{q}^*$ (the analogue of the complex conjugate) as

$$\mathbf{q} = w\mathbb{1} + x\mathsf{I} + y\mathsf{J} + z\mathsf{K}, \tag{1.47}$$

$$\mathbf{q}^* = w\mathbb{1} - x\mathsf{I} - y\mathsf{J} - z\mathsf{K}. \tag{1.48}$$

In the following, we will drop $\mathbb{1}$ when it is not necessary. With the definitions (1.44)–(1.47), we see that quaternion multiplication is not commutative, i.e. for two quaternions $\mathbf{q}_1$ and $\mathbf{q}_2$,

$$\mathbf{q}_1 \cdot \mathbf{q}_2 \neq \mathbf{q}_2 \cdot \mathbf{q}_1$$

in general. Nevertheless, the quaternion product is associative:

$$(\mathbf{q}_1 \cdot \mathbf{q}_2) \cdot \mathbf{q}_3 = \mathbf{q}_1 \cdot (\mathbf{q}_2 \cdot \mathbf{q}_3).$$

The dot '·' for quaternion multiplication is often omitted. Rules for quaternion conjugation are

$$(\mathbf{q}^*)^* = \mathbf{q}, \qquad (\mathbf{q}_1 + \mathbf{q}_2)^* = \mathbf{q}_1^* + \mathbf{q}_2^* \tag{1.49}$$

and, similar to matrix transposition,

$$(\mathbf{q}_1 \mathbf{q}_2)^* = \mathbf{q}_2^* \mathbf{q}_1^*. \tag{1.50}$$

Just as for complex numbers, the absolute value $|\mathbf{q}|$ of a quaternion can be defined via the quaternion's conjugate $\mathbf{q}^*$ as

$$|\mathbf{q}| = \sqrt{\mathbf{q} \cdot \mathbf{q}^*} = \sqrt{w^2 + x^2 + y^2 + z^2}.$$

A unit quaternion will be denoted by q and has an absolute value of 1:

$$|\mathsf{q}| = \sqrt{\mathsf{q}\mathsf{q}^*} = \sqrt{\mathsf{q}^*\mathsf{q}} = 1. \tag{1.51}$$

So $\mathbb{1}, \mathsf{I}, \mathsf{J}$ and K are all unit quaternions. Multiplication of a quaternion by a unit quaternion conserves the length of the original quaternion.

In the literature, different kinds of notation for quaternions are used. For example, the w from (1.47) is sometimes called the 'scalar part' s, while the triple $[x, y, z]$ (written with or without commas) may be called the 'vector part' $\mathbf{v}$.[3] Thus, the linear combination expression for a quaternion in (1.47) can be written in the following equivalent forms:

$$\mathbf{q} = w + x\mathsf{I} + y\mathsf{J} + z\mathsf{K} \tag{1.52}$$

$$= [x \quad y \quad z \quad w] \tag{1.53}$$

$$= (s, \mathbf{v}) \tag{1.54}$$

$$= [\mathbf{v}, s]. \tag{1.55}$$

Besides the component representation in (1.52), one might see tuples of coefficients as in (1.53), or pairs consisting of the scalar and vector parts as in (1.54) and (1.55). Most texts use round brackets if the vector part is written after the scalar part, as in (1.54), and square brackets when the vector part is written in front, as in (1.55). In contrast to linear algebra, it makes no sense to distinguish between row and column vectors in quaternion notation. The product of two quaternions $\mathbf{q}_1 = (s_1, \mathbf{v}_1)$ and $\mathbf{q}_2 = (s_2, \mathbf{v}_2)$ in scalar–vector notation turns out to be

$$\mathbf{q}_1 \cdot \mathbf{q}_2 = (s_1 s_2 - \mathbf{v}_1 \cdot \mathbf{v}_2, \; s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2), \tag{1.56}$$

with the usual cross product $\times$ and inner (dot) product $\cdot$ for three-dimensional vectors. Due to the anti-commutativity of the cross product ($\mathbf{v}_1 \times \mathbf{v}_2 = -\mathbf{v}_2 \times \mathbf{v}_2$), the quaternion product cannot be commutative; on the other hand, the inner product is not anti-commutative, so in general one must assume that $\mathbf{q}_1 \cdot \mathbf{q}_2 \neq \mathbf{q}_2 \cdot \mathbf{q}_1$. Quaternions do commute with their own conjugates, as the result is a scalar anyway, $\mathbf{q} \cdot \mathbf{q}^* = \mathbf{q}^* \cdot \mathbf{q} = |\mathbf{q}|^2$. For 'pure vectors' $\mathbf{r}_1 = (0, \mathbf{v}_1)$ and $\mathbf{r}_2 = (0, \mathbf{v}_2)$, we have, in the notation of (1.52)–(1.55),

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = (-\mathbf{v}_1 \cdot \mathbf{v}_2, \; \mathbf{v}_1 \times \mathbf{v}_2),$$

so their quaternion product contains the *negative* inner product in the scalar part and the cross product in the vector part.

---

[3]The conflict that arises with three-dimensional velocity vectors in using the symbol '$\mathbf{v}$' is unfortunate, but choosing a letter other than that used in the majority of texts on quaternions would not improve readability either; use of the $\mathbf{v}$ notation will be limited to this section and the next, where no rectilinear velocities occur.

In component notation, let us write

$$\mathbf{q}_1 = \begin{pmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{pmatrix}, \quad \mathbf{q}_2 = \begin{pmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix}, \quad \mathbf{q}_3 = \begin{pmatrix} w_3 \\ x_3 \\ y_3 \\ z_3 \end{pmatrix}.$$

Then, another way to write the quaternion product $\mathbf{q}_3 = \mathbf{q}_1 \cdot \mathbf{q}_2$ is

$$\begin{pmatrix} w_3 \\ x_3 \\ y_3 \\ z_3 \end{pmatrix} = \begin{pmatrix} w_1 & -x_1 & -y_1 & -z_1 \\ x_1 & w_1 & -z_1 & y_1 \\ y_1 & z_1 & w_1 & -x_1 \\ z_1 & -y_1 & x_1 & w_1 \end{pmatrix} \begin{pmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix}. \tag{1.57}$$

Inverses can be defined in the same way as for complex numbers, via the conjugate $\mathbf{q}^*$:

$$\mathbf{q}^{-1} = \frac{1}{|\mathbf{q}|} \mathbf{q}^* = \frac{1}{\sqrt{\mathbf{q} \cdot \mathbf{q}^*}} \mathbf{q}^*.$$

For unit quaternions with $|\mathbf{q}| = 1$, we have $\mathbf{q} \cdot \mathbf{q}^* = 1$ so that $\mathbf{q}^{-1} = \mathbf{q}^*$; hence the inverse of a unit quaternion is its conjugate. For an angle ϕ and a vector $\hat{\mathbf{v}}$, we can define a unit quaternion (sometimes called an 'Euler parameter')

$$\mathsf{q}_\phi = (\cos\phi, \ \hat{\mathbf{v}}\sin\phi), \tag{1.58}$$

which already looks a lot like the complex exponential we used in (1.43) to represent a rotation around the z-axis. The inverse of q_ϕ can then be written as

$$\mathsf{q}_\phi^{-1} = \mathsf{q}_\phi^* = (\cos\phi, \ \hat{\mathbf{v}}\sin\phi)^* = (\cos\phi, \ -\hat{\mathbf{v}}\sin\phi).$$

Next, we show how to represent vectors as quaternions. A 'pure' vector is a quaternion for which the scalar part is absent, and can be written in various forms as

$$\begin{aligned} \mathbf{r} &= 0 + x\mathsf{I} + y\mathsf{J} + z\mathsf{K} \\ &= [x \ \ y \ \ z \ \ 0] \\ &= (0, \mathbf{v}) \\ &= [\mathbf{v}, 0]. \end{aligned}$$

This means that the quaternion conjugate of a pure vector is the negative of the vector,

$$\mathbf{r}^* = -\mathbf{r}. \tag{1.59}$$

Now we figure out how to use quaternions to represent rotations of vectors. For complex numbers, rotation by angle ϕ is equivalent to multiplication by $\exp(\mathrm{i}\phi) = \cos\phi + \mathrm{i}\sin\phi$, so

let's see how far we can get with multiplying a pure vector quaternion $\mathbf{r} = x\mathsf{I} + y\mathsf{J} + z\mathsf{K}$ by a unit quaternion (Euler parameter) $\mathsf{q} = \cos\phi + \mathsf{I}\sin\phi$:

$$\begin{aligned}
\mathsf{q}\mathbf{r} &= (\cos\phi + \mathsf{I}\sin\phi)(x\mathsf{I} + y\mathsf{J} + z\mathsf{K}) \\
&= (x\mathsf{I} + y\mathsf{J} + z\mathsf{K})\cos\phi - x\sin\phi + y\sin\phi\mathsf{K} - z\sin\phi\mathsf{J} \\
&= x(\mathsf{I}\cos\phi - \sin\phi) + y(\mathsf{J}\cos\phi + \sin\phi\mathsf{K}) + z(\mathsf{K}\cos\phi - \sin\phi\mathsf{J}). \quad (1.60)
\end{aligned}$$

A rotation around the I-axis by q_I should leave the terms with $x\mathsf{I}$ in $\mathbf{v}$ unchanged, but this is not the case in (1.60). We can remedy this by multiplying (1.60) by q^* from the right, which gives

$$\begin{aligned}
\mathsf{q}\mathbf{r}\mathsf{q}^* &= (\cos\phi + \mathsf{I}\sin\phi)(x\mathsf{I} + y\mathsf{J} + z\mathsf{K})(\cos\phi - \mathsf{I}\sin\phi) \\
&= x\mathsf{I} + (y\cos 2\phi - z\sin 2\phi)\mathsf{J} + (y\sin 2\phi + z\cos 2\phi)\mathsf{K}. \quad (1.61)
\end{aligned}$$

Obviously, multiplication of $\mathbf{r}$ by q from the left and by q^* from the right is a rotation, as the length of $\mathbf{r}$ as well as the component in the direction of the rotation remain unchanged. Since the argument in (1.61) is 2ϕ, we have actually rotated $\mathbf{r}$ by 2ϕ. So, in order to define a quaternion associated with a rotation by angle ϕ, we need to halve the angle, i.e. use

$$\mathsf{q}_{\phi/2} = \cos\frac{\phi}{2} - \mathsf{I}\sin\frac{\phi}{2}. \quad (1.62)$$

The rotated vector $\tilde{\mathbf{r}}$ is then obtained by

$$\tilde{\mathbf{r}} = \mathsf{q}_{\phi/2}\mathbf{r}\mathsf{q}^*_{\phi/2}. \quad (1.63)$$

Therefore, although rotations around the z-axis for complex numbers (or two-dimensional real vectors) are expressed using a single complex multiplication by $\exp(i\phi)$, a general rotation in three dimensions requires multiplication by an Euler parameter $\mathsf{q}_{\phi/2}$ from the left and multiplication by the conjugate parameter $\mathsf{q}^*_{\phi/2}$ from the right. In general, the unit quaternion

$$\mathsf{q} = \left(\cos\frac{\phi}{2},\ \mathbf{v}\sin\frac{\phi}{2}\right) \quad (1.64)$$

represents a rotation around the unit vector $\mathbf{v}$; see Figure 1.13(b). General coordinate transformations for rotations of pure vectors $\mathbf{r}$ with quaternions q are obtained by

$$\tilde{\mathbf{r}} = \mathsf{q}\mathbf{r}\mathsf{q}^*. \quad (1.65)$$

Successive rotations, first by q and then by $\tilde{\mathsf{q}}$, have to be composed in the same way:

$$\tilde{\tilde{\mathbf{r}}} = \tilde{\mathsf{q}}\mathsf{q}\tilde{\mathbf{r}}\mathsf{q}^*\tilde{\mathsf{q}}^*.$$

From the point of view of the number of degrees of freedom, the use of unit quaternions to represent rotations means that instead of the 4-tuple of numbers $(s, \mathbf{v})$ for general quaternion

multiplications, we have an additional constraint, namely the length of the quaternion: only unit quaternions correspond to exactly three degrees of freedom for the three-dimensional rotation. Additionally, using Equations (1.59) and (1.50), one can show that

$$\tilde{\mathbf{r}} = -\left(\tilde{\mathbf{r}}\right)^* = -(\mathbf{q}(-\mathbf{r}^*)\mathbf{q}^*)^*$$
$$= ((-\mathbf{q})\mathbf{r}^*(-\mathbf{q}^*))^*$$
$$= (-\mathbf{q})\mathbf{r}(-\mathbf{q})^*. \tag{1.66}$$

This means that rotations can be represented either by quaternions q, as in (1.63), or by −q, as in (1.66). For the equations of motion, we also need the time derivatives of the quaternions and their relationship to the angular velocity $\boldsymbol{\omega}$. We list these formulae together here; their derivation will be the subject of the next subsection. The first derivative with respect to time of a unit quaternion q due to a rotation with angular velocity $\boldsymbol{\omega}$ is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q} = \dot{\mathbf{q}} = \frac{1}{2}\boldsymbol{\omega}(t)\,\mathbf{q}. \tag{1.67}$$

The second derivative, which is needed to write equations of motion in second order, like Newton's equation of motion, is

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathbf{q} = \ddot{\mathbf{q}} = \frac{1}{2}\left(\dot{\boldsymbol{\omega}}\mathbf{q} + \dot{\mathbf{q}}\boldsymbol{\omega}\right). \tag{1.68}$$

Additionally, we will need the following auxiliary equations:

$$\boldsymbol{\omega} = 2\dot{\mathbf{q}}\mathbf{q}^*, \tag{1.69}$$

$$\dot{\boldsymbol{\omega}} = \mathbf{J}^{-1}\left(\mathbf{L} \times \boldsymbol{\omega} + \boldsymbol{\tau}\right), \tag{1.70}$$

$$\mathbf{L} = \mathbf{J}\boldsymbol{\omega}, \tag{1.71}$$

$$\mathbf{J} = \mathbf{A}\mathbf{J}^{\mathrm{b}}\mathbf{A}^{\mathrm{T}}, \tag{1.72}$$

$$\mathbf{J}^{-1} = \mathbf{A}(\mathbf{J}^{\mathrm{b}})^{-1}\mathbf{A}^{\mathrm{T}}, \tag{1.73}$$

where the angular momentum **L**, angular velocity $\boldsymbol{\omega}$, torque $\boldsymbol{\tau}$, and the quaternion q and its time derivatives are computed in the space-fixed coordinate system. The moment of inertia **J** is also calculated in the space-fixed coordinate system, which can be obtained from the moment of inertia in the body-fixed system $\mathbf{J}^{\mathrm{b}}$ via multiplication by the rotation matrix **A** from the left and its inverse $\mathbf{A}^{\mathrm{T}}$ from the right (the 'principal axis transform').

At the initialization of a DEM simulation, the original orientation of a particle is chosen in the body-fixed coordinate system where its tensor of inertia is diagonal, $\mathbf{J} = (J_x, J_y, J_z)$. Each particle is then rotated into the space-fixed (i.e. Cartesian, axis-aligned) coordinate system using the value of the orientation quaternion $\mathbf{q}(t = 0)$ which uniquely defines the orientation of each particle at each time-step t. The transformation of the tensor of inertia is better made with the rotation matrix **A** than with the quaternion of orientation. In the next subsection we give a representation of **A** in components of q. The equations of motion are second-order differential equations of unit quaternions, (1.68) in the body-fixed coordinate system, and can be

integrated directly by the Gear predictor–corrector method, as no additional transformations are necessary. Other representations are possible. Allen and Tildesley [4] use unit quaternions q^b together with angular velocities $\boldsymbol{\omega}^b$ in the body-fixed coordinate system for first-order equations of angular motion. In their approach, torques $\boldsymbol{\tau}^b$ are computed in the space-fixed coordinate system as $\boldsymbol{\tau}^s$ and then transformed to the body-fixed coordinate system. However, for quaternions q^b in the body-fixed coordinate system, the higher-order derivatives needed for predictor–corrector time integration schemes (in particular the backward difference formulae we will recommend later) are not available, so other time integrators have to be used, which may not have the same favorable stability properties.

1.3.10 Derivation of quaternion dynamics

In this subsection, we give several relations for unit quaternions q and their time derivatives, and show derivations where comparing different references may lead to confusion about the meaning of the equations. The unit quaternion $q = (q_0, q_1, q_2, q_3)$ has length 1; its squared absolute value

$$|qq| = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \tag{1.74}$$

is a constant, so its time derivative (computed componentwise) must vanish:

$$\frac{\mathrm{d}}{\mathrm{d}t}|qq| = q_0\dot{q}_0 + q_1\dot{q}_1 + q_2\dot{q}_2 + q_3\dot{q}_3 = 0. \tag{1.75}$$

The middle part of (1.75) is a scalar product of the quaternion q with its own time derivative $\dot{q}$. This means that q is orthogonal to its own time derivative $\dot{q}$, similar to Figure 1.11, and similar to the orthogonality in the constraint coordinates for the pendulum discussed in § 2.8. Moreover, (1.75) implies that a quaternion and its time derivative commute: $\dot{q}q = q\dot{q}$. The components of q can be expressed by the Euler angles θ, ψ, ϕ from § 1.3.4 as follows:

$$q_0 = \cos\frac{\theta}{2}\cos\frac{\phi + \psi}{2}, \tag{1.76}$$

$$q_1 = \sin\frac{\theta}{2}\cos\frac{\phi - \psi}{2}, \tag{1.77}$$

$$q_2 = \sin\frac{\theta}{2}\sin\frac{\phi - \psi}{2}, \tag{1.78}$$

$$q_3 = \cos\frac{\theta}{2}\sin\frac{\phi + \psi}{2}. \tag{1.79}$$

This means that the components of the unit quaternion q can represent the orientation of a particle without involving trigonometric functions of the Euler angles. As the numerical evaluation of trigonometric functions takes about ten times longer than multiplication or addition, this means that unit quaternions are also computationally economical. With (1.76)–(1.79), we can express the rotation matrix as

$$\mathbf{A} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 + q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}.$$

Next, we compute the time derivative $\dot{q}$. Consider an infinitesimal rotation by a very small angle θ around a vector $\mathbf{v}$, so that the corresponding unit quaternion is

$$\left(\cos \frac{\theta}{2}, \ \mathbf{v} \sin \frac{\theta}{2} \right) \approx \left(1, \mathbf{v} \frac{\mathrm{d}\theta}{2} \right). \tag{1.80}$$

Let the rotation take place during the infinitesimal time interval $\mathrm{d}t$. The rotation experienced by a pure vector at time t, $\mathbf{r}(t) = (0, \mathbf{r})$, is then (see [8])

$$\mathbf{r}(t + \mathrm{d}t) = \mathbf{r} + \mathrm{d}\mathbf{r} = \left(1, \ \mathbf{v} \frac{\mathrm{d}\theta}{2} \right) (0, \mathbf{r}) \left(1, \ - \mathbf{v} \frac{\mathrm{d}\theta}{2} \right) \tag{1.81}$$

$$= \left(\mathbf{v} \frac{\mathrm{d}\theta}{2} \cdot \mathbf{r} - \mathbf{r} \left(-\mathbf{v} \frac{\mathrm{d}\theta}{2} \right), \ \mathbf{r} + \mathbf{v} \frac{\mathrm{d}\theta}{2} \times \mathbf{r} + \mathbf{r} \left(-\mathbf{v} \frac{\mathrm{d}\theta}{2} \right) \right) \tag{1.82}$$

$$= (0, \ \mathbf{r} + \mathrm{d}\theta \, (\mathbf{v} \times \mathbf{r})). \tag{1.83}$$

From (1.81) to (1.82), we multiplied out all terms and dropped the ones which are quadratic in $\mathrm{d}\theta$; from (1.82) to (1.83), we used the fact that the vector cross product is anti-commutative, $\mathbf{r} \times \mathbf{v} = -\mathbf{v} \times \mathbf{r}$. It follows that for the motion from $\mathbf{r}(t)$ to $\mathbf{r}(t + \mathrm{d}t)$,

$$\mathbf{r}(t + \mathrm{d}t) - \mathbf{r}(t) = \mathrm{d}\mathbf{r} = \mathrm{d}\theta \, (\mathbf{v} \times \mathbf{r}). \tag{1.84}$$

If we divide (1.84) by $\mathrm{d}t$, we obtain

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \dot{\mathbf{r}} = \frac{\mathrm{d}\theta}{\mathrm{d}t} (\mathbf{v} \times \mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}, \tag{1.85}$$

because the angular velocity $\boldsymbol{\omega}$ in the space-fixed system is around the direction $\mathbf{v}$. Next we want to derive the relationship between q and the angular velocity $\boldsymbol{\omega}^b$ in the body-fixed coordinate system. We cannot obtain $\boldsymbol{\omega}^b$ by simple quaternion transformations as in Equation (1.85), because we have $\boldsymbol{\omega}$ only as a factor in a vector product, not isolated on one side of an equation. Let us consider the time derivative of $\mathbf{r}$, treated as the orientation $\mathbf{r}^b$ of the body-fixed coordinate system rotated by a unit quaternion q. From the coordinate transformation in (1.65), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{r} = \frac{\mathrm{d}}{\mathrm{d}t} \left(q \mathbf{r}^b q^* \right)$$

$$= \left(\frac{\mathrm{d}}{\mathrm{d}t} q \right) \left(\mathbf{r}^b q^* \right) + q \underbrace{\left(\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{r}^b \right)}_{0} q^* + \left(q \mathbf{r}^b \right) \left(\frac{\mathrm{d}}{\mathrm{d}t} q^* \right)$$

$$= \left(\frac{\mathrm{d}}{\mathrm{d}t} q \right) \left(\mathbf{r}^b q^* \right) + \left(q \mathbf{r}^b \right) \left(\frac{\mathrm{d}}{\mathrm{d}t} q^* \right) \tag{1.86}$$

where $\mathbf{r}^b$ is fixed, so its time derivative vanishes. Replacing $\mathbf{r}^b$ with the transformed $\mathbf{r}$, i.e. substituting

$$\mathbf{r}^b = q^* \mathbf{r} q$$

into Equation (1.86), gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r} = \left(\frac{\mathrm{d}}{\mathrm{d}t}q\right)\left(\mathbf{r}^b q^*\right) + \left(q\mathbf{r}^b\right)\left(\frac{\mathrm{d}}{\mathrm{d}t}q^*\right)$$

$$= (\dot{q})\left(q^*\mathbf{r}qq^*\right) + \left(qq^*\mathbf{r}q\right)(\dot{q}^*)$$

$$= \dot{q}q^*\mathbf{r} - \mathbf{r}\dot{q}^*q. \tag{1.87}$$

We use that $\dot{q}^*$ and q commute. The vectorial components $(\dot{q}q^*)^{\mathrm{vec}}$ of $\dot{q}q^*$ (the components of the 'pure vector') will behave like vectors in the multiplication with $\mathbf{r}$, i.e. the product is anti-commutative:

$$(\dot{q}q^*)^{\mathrm{vec}}\mathbf{r} = -\mathbf{r}(\dot{q}q^*)^{\mathrm{vec}}.$$

Accordingly, we can transform the quaternion Equation (1.87) into a vector equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r} = (\dot{q}q^*)^{\mathrm{vec}} \times \mathbf{r} - \mathbf{r} \times (\dot{q}^*q)^{\mathrm{vec}}$$

$$= (\dot{q}q^*)^{\mathrm{vec}} \times \mathbf{r} + (\dot{q}^*q)^{\mathrm{vec}} \times \mathbf{r}$$

$$= 2(\dot{q}q^*)^{\mathrm{vec}} \times \mathbf{r}. \tag{1.88}$$

Comparing this with Equation (1.85), we see that

$$2(\dot{q}q^*)^{\mathrm{vec}} = \boldsymbol{\omega}. \tag{1.89}$$

Now we multiply (1.89) from the left by q^* and from the right by q to obtain

$$q^*2(\dot{q}q^*)^{\mathrm{vec}}q = \underbrace{q^*\boldsymbol{\omega}q}_{\boldsymbol{\omega}^b},$$

$$q^*2\dot{q} = \boldsymbol{\omega}^b, \tag{1.90}$$

so the right-hand side is the body-fixed angular velocity $\boldsymbol{\omega}^b$. Multiplying (1.90) from the left by q gives

$$qq^*2\dot{q} = q\boldsymbol{\omega}^b,$$

$$\dot{q} = \frac{1}{2}q\boldsymbol{\omega}^b. \tag{1.91}$$

This result is not a contradiction to (1.87), which involved the angular frequency from the space-fixed coordinate system, $\boldsymbol{\omega}$. The most important relations for quaternions are summarized in Table 1.2.

1.4   The phase space

In the phase space concept, the position coordinates $x(t)$ and the velocities $v(t)$ are considered together when analyzing a mechanical system. The mathematical theory of phase space analysis for ordinary differential equations is often called the theory of dynamical systems, and is

Table 1.2   Summary of important quaternion relations.

Coordinate representation of quaternions	$\mathbf{q} = w\mathbb{1} + x\mathsf{I} + y\mathsf{J} + z\mathsf{K}$ $= [x \quad y \quad z \quad w]$ $= (s, \mathbf{v})$ $= [\mathbf{v}, s]$	Pure vectors represented as quaternions	$\mathbf{r} = 0 + x\mathsf{I} + y\mathsf{J} + z\mathsf{K}$ $= [x \ y \ z \ 0]$ $= (0, \mathbf{v})$ $= [\mathbf{v}, 0]$
Conjugate quaternion	$\mathbf{q}^* = w\mathbb{1} - x\mathsf{I} - y\mathsf{J} - z\mathsf{K}$		
Rules for quaternion conjugation	$(\mathbf{q}^*)^* = \mathbf{q}$ $(\mathbf{q}_1 + \mathbf{q}_2)^* = \mathbf{q}_1^* + \mathbf{q}_2^*$	Euler parameter for vector $\mathbf{v}$ and angle ϕ, and its inverse	$\mathsf{q}_\phi = (\cos\phi, \hat{\mathbf{v}}\sin\phi)$ $\mathsf{q}_\phi^{-1} = \mathsf{q}_\phi^* = \mathsf{q}_{-\phi}$ $= (\cos\phi, \hat{\mathbf{v}}\sin\phi)^*$ $= (\cos\phi, -\hat{\mathbf{v}}\sin\phi)$
Inverse quaternion	$\mathbf{q}^{-1} = \dfrac{1}{|\mathbf{q}|}\mathbf{q}^* = \dfrac{1}{\sqrt{\mathbf{q}\cdot\mathbf{q}^*}}\mathbf{q}^*$	Rotation of a pure vector by angle ϕ	$\tilde{\mathbf{r}} = \mathsf{q}_{\frac{\phi}{2}}\,\mathbf{r}\,\mathsf{q}_{\frac{\phi}{2}}^*$
Quaternion product	$\mathbf{q}_1 \cdot \mathbf{q}_2 = (s_1 s_2 - \mathbf{v}_1\cdot\mathbf{v}_2, \\ s_1\mathbf{v}_2 + s_2\mathbf{v}_1 + \mathbf{v}_1\times\mathbf{v}_2)$		
No commutativity in general	$\mathbf{q}_1 \cdot \mathbf{q}_2 \neq \mathbf{q}_2 \cdot \mathbf{q}_1$	Quaternion time derivatives and angular velocity ($\boldsymbol{\omega}$ in space-fixed, $\boldsymbol{\omega}^{\mathrm{b}}$ in body-fixed coordinates)	$\dfrac{\mathrm{d}}{\mathrm{d}t}\mathsf{q} = \dot{\mathsf{q}} = \dfrac{1}{2}\boldsymbol{\omega}(t)\,\mathsf{q}$ $\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}\mathsf{q} = \ddot{\mathsf{q}}$ $= \dfrac{1}{2}(\dot{\boldsymbol{\omega}}\mathsf{q} + \dot{\mathsf{q}}\boldsymbol{\omega})$ $\boldsymbol{\omega} = 2\dot{\mathsf{q}}\mathsf{q}^*$ $\dot{\mathsf{q}} = \dfrac{1}{2}\mathsf{q}\boldsymbol{\omega}^{\mathrm{b}}$
Associativity	$(\mathbf{q}_1 \cdot \mathbf{q}_2) \cdot \mathbf{q}_3 = \mathbf{q}_1 \cdot (\mathbf{q}_2 \cdot \mathbf{q}_3)$		
Conjugate of a product	$(\mathbf{q}_1\mathbf{q}_2)^* = \mathbf{q}_2^*\mathbf{q}_1^*$		
Absolute value	$|\mathbf{q}| = \sqrt{\mathbf{q}\cdot\mathbf{q}^*} = \\ \sqrt{w^2 + x^2 + y^2 + z^2}$		
Unit quaternion	$\mathsf{q} = \dfrac{1}{\sqrt{\mathbf{q}\mathbf{q}^*}}\mathbf{q}$ $|\mathsf{q}| = \sqrt{\mathsf{q}\mathsf{q}^*} = \sqrt{\mathsf{q}^*\mathsf{q}} = 1$		

a generalization of the theory of ordinary differential equations. We will use the terms 'flow of the differential equation', 'dynamical system' and 'phase space' pretty synonymously. In physics, a phase space is usually the space spanned by the coordinates and their respective velocities or, in some cases, momenta. In mathematics, any pair of coordinates may be considered for a dynamical system. This implies a certain symmetry between coordinates and their velocities, which is not present in the more elementary kinematics, where velocities are 'merely' the time derivatives of the original coordinates. The main types of flow that are of
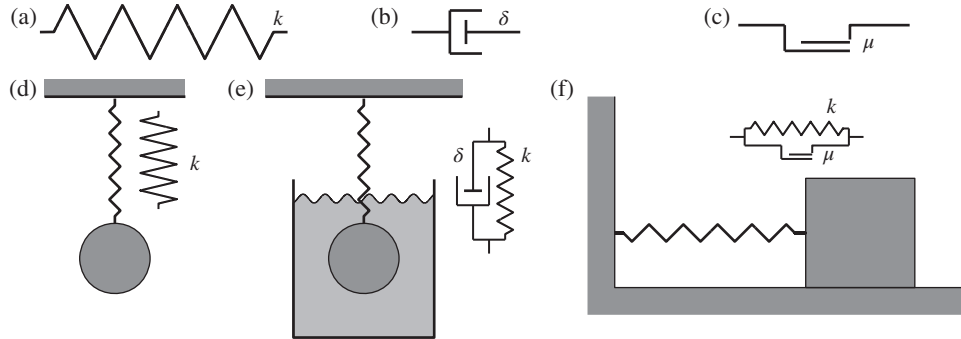
Figure 1.14 Symbols used in mechanical engineering: (a) a spring with spring constant k; (b) viscous (velocity-dependent) damping (dashpot) proportional to $\delta\dot{x}$; (c) system with (Coulomb) friction (slider) proportional to μ; (d) mass dangling from a spring, the physical system corresponding to (a); (e) mass dangling from a spring in water (viscous damping, effectively spring–dashpot), the physical system corresponding to (b); (f) block sliding on a surface under the influence of a spring (slider–spring), the physical system corresponding to (c). In (d)–(f) the symbols are shown as insets. Discrete element interactions are sometimes sketched with a spring–dashpot element in the normal direction and a slider–spring symbol in the tangential direction.

interest for discrete element systems will be discussed in the context of the linear oscillators example (with elastic force proportional to the dislocation), for both undamped and damped cases as well as for the case with Coulomb friction; for the symbols used in engineering and illustrations of the corresponding physical systems, see Figure 1.14. The flow in phase space will enable us to discriminate between static and dynamic friction, even in the case where the sliding velocity is zero.

### 1.4.1 Qualitative discussion of the time dependence of linear oscillations

Here we give a brief hand-waving derivation of the phase flow as a basis for discussing the dynamics in more depth later. Detailed time-dependent solutions for the linear oscillator can be found, for instance, in Benenson [9]. The undamped linear oscillator, or 'harmonic oscillator', corresponds to an ideal elastic (meaning there is no dissipation) linear spring modeled by the differential equation

$$m\ddot{x} + kx = 0. \tag{1.92}$$

Further, it corresponds to the motion of a particle in a symmetric force equilibrium. One solution of (1.92) is $x(t) = A\cos(\omega_0 t + \phi_0)$, where $\omega_0 = \sqrt{k/m}$. Apart from the phase parameter ϕ_0, the corresponding velocities are $v = \dot{x}(t) = -A\omega_0 \sin\omega t$. Neglecting the constant pre-factors, which is equivalent to a simple rescaling of the axis, we have

$$x \propto \cos(\omega_0 t), \tag{1.93}$$

$$v \propto \sin(\omega_0 t), \tag{1.94}$$
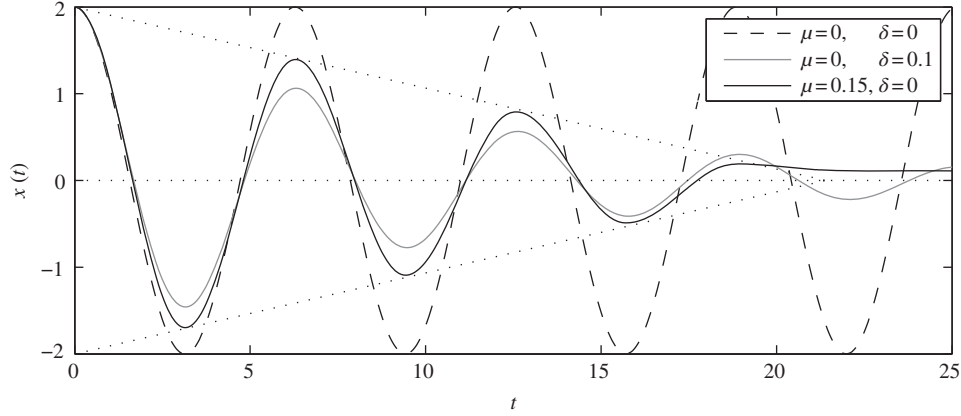
Figure 1.15 Graph of the amplitude of the linear oscillator with $m = 1$ and $k = 1$: damped (dashed line); viscously damped (solid gray line); and damped with Coulomb friction (solid black line), together with the envelope for the extrema of the graph in the Coulomb friction case and the zero-amplitude x-axis (three dotted lines).
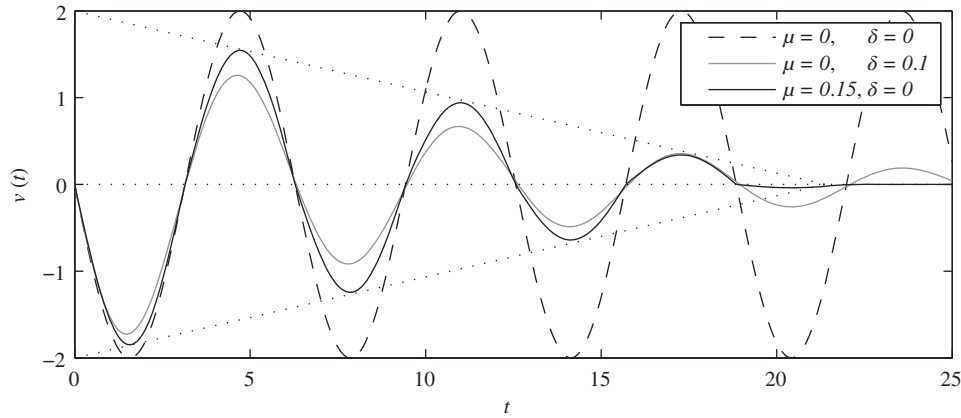
Figure 1.16 Graph of the velocity of the linear oscillator with $m = 1$ and $k = 1$: damped (dashed line); viscously damped (solid gray line); and damped with Coulomb friction (solid black line), together with the envelope for the extrema of the graph in the Coulomb friction case and the zero-velocity x-axis (three dotted lines).

i.e. the position (dashed line in Figure 1.15) and velocity (dashed line in Figure 1.16) variables oscillate with a phase difference of $\pi/2$ between them.

For the linear oscillator with viscous damping (proportional to velocity), the differential equation is

$$m\ddot{x} + 2\delta\dot{x} + kx = 0. \tag{1.95}$$

Physically, it corresponds to a linear spring which is damped in a fluid, so that the damping stays proportional to the velocity. If there were inertia effects in the fluid, the resulting

'Newtonian friction' would be proportional to the square of the velocity. The solution to the viscously damped linear oscillator equation is

$$x(t) = A \exp(-\delta t) \exp\left(\pm i\sqrt{\omega_0^2 - \delta^2}t\right) \quad \text{where } \omega_0 = \sqrt{\frac{k}{m}}, \ \delta = \frac{b}{2m}. \tag{1.96}$$

We focus on oscillatory solutions (with $\omega_0^2 > \delta^2 t$); see the solid gray curve in Figure 1.15. The critically damped and over-damped cases can be found in Benenson [9]. The exponential decay of the solution, $x(t) \propto \exp(-\delta t)$, leads to a similar exponential decay of the velocities (see Figure 1.16, solid gray line). Continuum materials under vibration usually show viscous damping patterns, too, due to the dissipation mechanisms of kinetic energies in solids. Exponential decay sounds impressive, but is in fact a relatively 'weak' type of decay: the amplitude never actually reaches zero.

With Coulomb friction (dry friction or sliding friction), the linear oscillator becomes

$$m\ddot{x} + \mu \, \mathfrak{sgn}(\dot{x}) + kx = 0, \tag{1.97}$$

where μ is the product of the friction coefficient and the normal force, and we define the $\mathfrak{sgn}$ function as

$$\mathfrak{sgn}(a) \begin{cases} = 1, & \text{for } a > 0, \\ \in [-1, 1] & \text{for } a = 0, \\ = -1 & \text{for } a < 0, \end{cases} \tag{1.98}$$

so that the friction force exactly compensates for the external force. Note that this is different from the usual step function definition

$$\text{sgn}(a) = \begin{cases} 1 & \text{for } a > 0, \\ 0 & \text{for } a = 0, \\ -1 & \text{for } a < 0. \end{cases}$$

Physically, the system corresponds to a spring that is fixed to a wall and connected to a block which slides on the floor nearby; see Figure 1.14(f). In this chapter the discussion will be in a hand-waving fashion; we give the exact solution for $v = 0$ in Chapter 3. For $\mathfrak{sgn}(v) = \mathfrak{sgn}(\dot{x}) = \pm 1$, the solution is composed of solutions to one of the inhomogeneous differential equations [10, 11]

$$m\ddot{x} + kx = -\mu, \tag{1.99}$$

$$m\ddot{x} + kx = +\mu, \tag{1.100}$$

or the amplitude stays constant when $-kx$ is smaller than μ. The solutions to Equations (1.99) and (1.100) have the same periodicity as the solution to (1.92), with $\omega_0 = \sqrt{k/m}$, because for a linear ordinary differential equation, introducing a non-zero term on the right-hand side (inhomogeneity) does not change the general solution. The effect of damping with Coulomb friction is that the piecewise solution branches between the reversals in sign of the

velocity decay in magnitude (for both the amplitude and the velocity) within a linear envelope (the outer dotted lines in Figures 1.15 and 1.16). This means that the relative maxima of the positions and velocities lie along a line, and likewise for the minima, so that after a finite time, the velocity $v(t)$ becomes zero and the amplitude $x(t)$ becomes constant. As can be seen in Figure 1.15, the final amplitude does not have to be zero: when the spring force $-kx$ is smaller than the friction force μ, the amplitude stays fixed, which is why we have to use the inclusion definition for the sign in (1.98). From Figure 1.16 one sees that the velocity, and with it the kinetic energy, goes to zero in finite time, so Coulomb friction is much more effective in damping out energies or vibrations than is velocity-dependent friction, especially at small velocities. This effect has various applications. Machine parts (e.g. running gears and wheels of trains) are tested by tapping them with a hammer. If everything is in good condition, one hears a nice 'metallic' ringing sound: the sound amplitude is damped out exponentially and decays smoothly. If there are cracks, the contact between ragged surfaces damps the sound much faster due to Coulomb friction, so that it comes out as a short, ugly rattling noise. One can visualize this effect by fixing one end of a ruler on a desk and setting the other end to vibrate; usually there will be a smooth decay in the vibration amplitude, but if the vibrating end is in frictional contact with another object, the decay will be abrupt.

Individually, contacts in granular assemblies are equivalent to linear oscillators with Coulomb friction. For this reason, aggregates of granular material are often much better at damping out kinetic energies than a similar piece of continuum material would be. Jugglers use grain-filled balls for practice, because such balls won't roll away when accidentally dropped; sand slopes are used in shooting ranges to catch straying bullets, while sand sacks are used for protection against aimed bullets.

1.4.2 Resonance

Now let us consider what happens when we drive the damped linear oscillator of Equation (1.95) by a periodic force which oscillates with period ω and maximal amplitude $\tilde{f}_0$. To reduce the amount of algebra required for the solution, we write the periodic force in complex exponential form, so that the equation is

$$\ddot{x} + 2\delta\dot{x} + \omega_0^2 x = \tilde{f}_0 \exp(i\omega t) \tag{1.101}$$

where $\omega_0 = \sqrt{k/m}$. We are interested in the absolute value of the amplitude A of the stationary solution

$$x(t) = A \exp(i\omega t). \tag{1.102}$$

Substituting (1.102) into (1.101) allows us to get rid of the time dependence (by canceling out factors of $\exp(i\omega t)$) and hence obtain[4]

$$-A\omega^2 + i2\delta A\omega + \omega_0^2 A = f_0, \tag{1.103}$$

---

[4]This is possible because of how we captured the time dependence with a complex exponential; to formulate a solution using only real functions, about two pages of arithmetic and algebraic transformations are necessary; see, for instance, Knudsen and Hjorth [12, § 15.6].
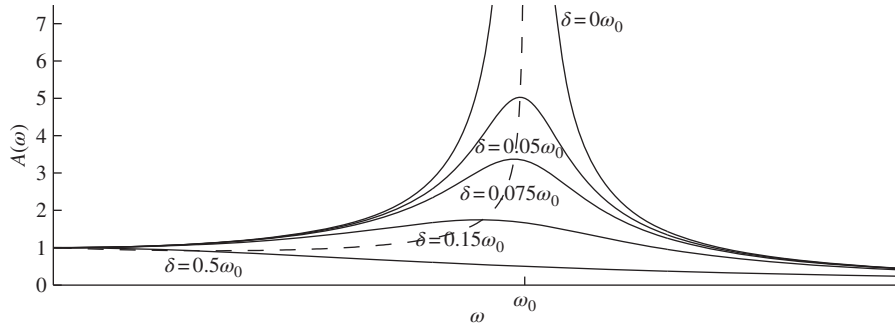
Figure 1.17 Graph of the absolute value of the resonance amplitude, $|A|$, as a function of ω for the linear oscillator in Equation (1.103) with various values of δ. The dashed line gives the positions of the maxima as in formula (1.106).

with $f_0 = \tilde{f}_0/m$, which then gives

$$A = \frac{f_0}{-\omega^2 + \mathrm{i}2\delta\omega + \omega_0^2}. \tag{1.104}$$

From this, we obtain the absolute value of the complex amplitude A according to (1.42):

$$|A| = \frac{f_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2}}. \tag{1.105}$$

(The absolute value is also more meaningful in the purely real case with $\delta = 0$, as $A(\omega)$ changes sign from $+\infty$ to $-\infty$ at $\omega = \omega_0$; since we are interested in the magnitude of the amplitude, the sign is not important.) The resulting amplitudes are plotted in Figure 1.17 for several values of δ. For damping $0 \leq \delta \leq 1$, the maxima of the resonance amplitudes lie on the curve

$$A_{\max} = \frac{f_0}{\delta\sqrt{(\omega_0^2 - \delta^2)}}. \tag{1.106}$$

For $\delta = 0$, the amplitude increases toward infinity, i.e. an undamped system excited at the resonance frequency $\omega = \omega_0$ would be destroyed, due to unlimited growth of the vibration amplitude. Note that the amplitude increases only linearly in time, so that an infinite amplitude would only be reached after an infinite amount of time; see Exercise 1.3. The right-hand side $f_0 \exp(\mathrm{i}\omega t)$ of Equation (1.101) contains only a t-dependence, so it is an 'external' force; terms with dependence on x only are the 'internal' forces of the system. In mathematical terminology, systems that depend only on 'x' are said to be autonomous, while those which also have a dependence on 't' are non-autonomous.

1.4.3 The flow in phase space

With the results from the previous subsection, we are ready to discuss the flow of the differential equation in phase space, also called the 'attractor' of the system. (The flow will be used

later in Chapter 3 to make mathematically exact distinctions between conditions for static and dynamic friction.) In Figures 1.18–1.20, we visualize the flow in several ways. First, we plot with solid lines the trajectories in time, $(x(t), v(t))$, of the solution. We can also consider Newton's equation of motion in the form (1.7)–(1.8), written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} v \\ F/m \end{pmatrix}, \tag{1.107}$$

so that the right-hand side is equivalent to the directions

$$\begin{pmatrix} \dfrac{x(t+\delta t)-x(t)}{\delta t} \\ \dfrac{v(t+\delta t)-v(t)}{\delta t} \end{pmatrix} = \begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} \tag{1.108}$$

of the flow field: these directions are depicted as arrows in Figures 1.18–1.20. Finally, it has become traditional to discuss the transport of a set of initial conditions in phase space from time t_0 to time t:

$$y(t_0) \rightarrow y(t). \tag{1.109}$$
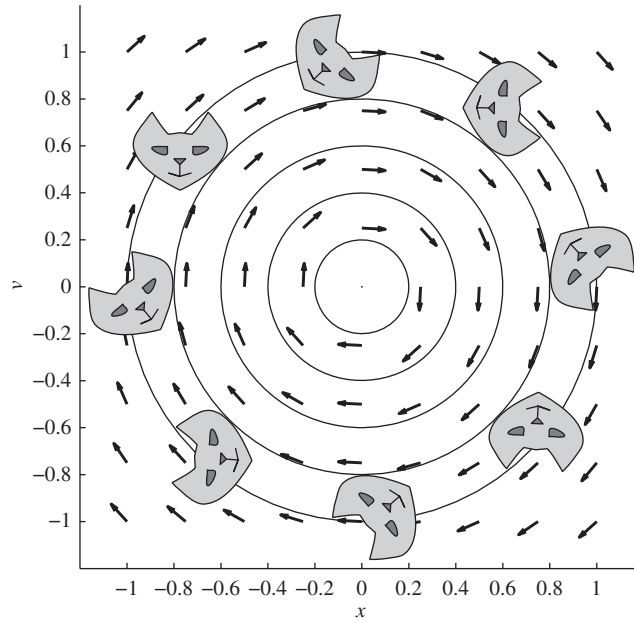
Figure 1.18 Phase portrait (attractor) for the linear oscillator ($m = 1$, $k = 1$) without damping: illustration of Liouville's theorem on conservation of phase space volume.
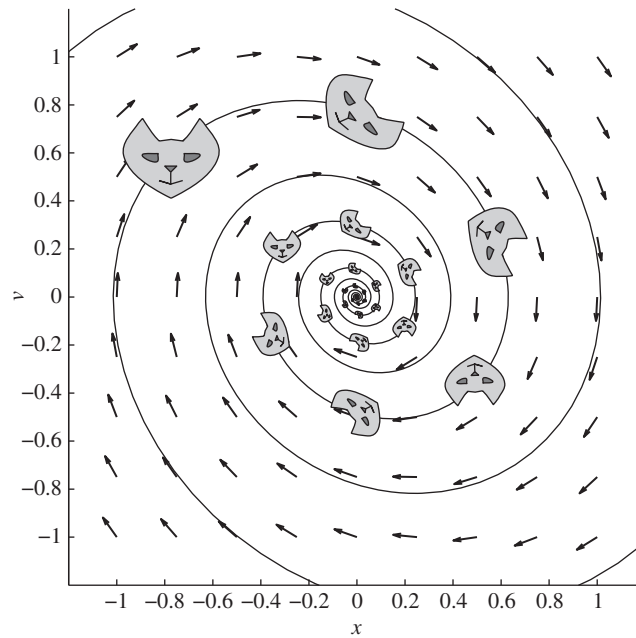
Figure 1.19 Phase portrait for the linear oscillator ($m = 1$, $k = 1$) with viscous damping ($\delta = 0.1$): the attractor is a whirl, where the phase space volume shrinks exponentially in spiral-shaped trajectories.
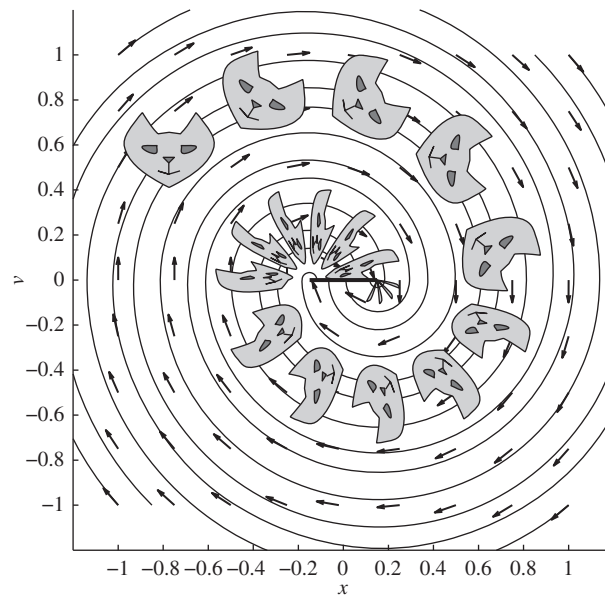
Figure 1.20 Attractors for the linear oscillator ($m = 1$, $k = 1$) with Coulomb friction ($\mu = 0.15$): the flow of an initially simply connected phase space volume is split and deposited from above and from below at the singularity $g(x, v) = (-\mu \leq x \leq \mu, \ v = 0)$.

The initial conditions usually take the shape of a cat's head, which goes back to Arnold's book on mechanics [13], but is probably originally due to Delaunay.[5]

For the undamped linear oscillator, plotting sine against cosine from the solutions (1.93)–(1.94) gives circular trajectories, as shown in Figure 1.18. We can see that the area of the cat's head does not change, i.e. it is a 'conserved' quantity; this illustrates Liouville's theorem, which says that phase space density is conserved for 'Hamiltonian' mechanical systems, i.e. systems for which Newton's equation of motion can be written as [15]

$$\frac{\mathrm{d}}{\mathrm{d}t}x = m^{-1}p,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p = -\nabla_x V(x),$$

where x denotes position, m is mass, p is momentum, and $\nabla_x V(x)$ is the gradient of the position-dependent potential $V(x)$. In mathematics, such systems of ordinary differential equations are said to be 'symplectic'; in physics they are called 'Hamiltonian' or 'canonical' systems [3, 13, 16]. Among other things, these systems exhibit conservation of energy. The direction field is continuous, i.e. the mapping

$$\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \mapsto \begin{pmatrix} x(t + \delta t) + \delta x \\ v(t + \delta t) + \delta v \end{pmatrix} \tag{1.110}$$

with infinitesimal δt, δx and δv is continuous for all initial values of $x, v > 0$.

When damping is introduced, the amplitude in (1.96) decays exponentially; see Figure 1.19. Viscous damping leads to an exponential contraction of the cat's head, i.e. the volume spanned by the initial condition decreases during transport of the coordinates in phase space, but the shape stays basically the same. The exponential decay gives spiral- or vortex-shaped trajectories in phase space, or whirls, as they are called in the field of dynamical systems. As for the energy-conserving system in (1.110), the right-hand side functions in Equation (1.107) are also continuous from one point to another in phase space, and the direction field has no singularity; in other words, the direction change from an arrow at $(x(t), v(t))$ to a nearby arrow at $(x(t + \delta t) + \delta x, v(t + \delta t) + \delta v)$ is always smooth, and the singularity $(v = 0, x = 0)$ cannot be reached in finite time, so it is not part of the phase space for the problem.

The situation changes dramatically when we have Coulomb friction; see Figure 1.20. At the beginning the attractor resembles that in the viscous damping case: for $(|\mu| > x, v = 0)$, the situation for dynamic friction, the flow is continuous. Along $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$, the flow is non-smooth. In an infinitesimal region around $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$, flow from above or from below can occur: this is the region of static friction, where the tension of the spring at finite displacement in Figure 1.14 is not strong enough to overcome the friction force acting on the block. When the cat's head approaches the line $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$, it splits up: part of the flow is transported into $g(x, v)$ from above, another part from below. No flow is possible on the horizontal axis, either from left to right or from right to left. This is a consequence of the fact that the right-hand side of the

---

[5]According to Zdravkovska et al. [14, p. 82], B. N. Delaunay (1890–1980), who taught at Moscow University where Arnold studied, used to visualize affine transformations by 'transformations of a picture of a kitten'.

system (1.99)–(1.100) is not smooth: arrows coming from below face upward, arrows coming from above face downward, and along the whole line $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$ arrows have zero length. Note that while across the line the flow is not smooth, the line itself is part of the phase space of the problem and corresponds to the situation in Figure 1.14(f), where the spring is under tension but the block does not move because it is held by the friction force. A contraction like for viscous damping has been proposed [17] for the phase space evolution of, among other systems, sheared granular materials, which implies a flow as in Figure 1.19. As these materials are, to all intents and purposes, assemblies of solid particles with Coulomb friction (except in the most artificial cases), assumption of a 'damped' Liouville equation

$$\frac{\partial f(t)}{\partial t} = -[i\mathcal{L} + \Lambda] f(t) = -i\tilde{\mathcal{L}} f(t) \tag{1.111}$$

to describe the phase space volume, with a solution that is an exponential contraction of the phase space $f(t) = \exp(-i\tilde{\mathcal{L}}t) f_0$ of Equation (1.111), may be an appropriate local description in some cases; however, globally this approach is inappropriate, even for only a single frictional contact, as a comparison of Figures 1.19 and 1.20 easily shows. Such a description would easily break down for transitions from dynamic to static friction, for example from hopper flow to clogging. This provides a more esoteric justification of why one should use particle models with Coulomb friction: they allow us to access much more exotic flows in phase space than do continuum approaches alone.

While physically $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$ is reached in finite time in Figure 1.20, Filippov theory [18], the standard theory for differential equations with discontinuous right-hand sides, does not allow for singularities $g(x, v) = (-\mu \leq x \leq \mu, v = 0)$ which have the shape of a line in the solution domain. Instead of the attractor in Figure 1.20, for singularities in the flow directions in (1.108), Filippov theory [18, Ch. 4] postulates transport along the line (in our case, along the x-axis with $v = 0$), but this is clearly impossible in the case of a spring with a block: the block can only change its position if its velocity is finite. If the singularities are reached only as $t \to \infty$, this may be physically meaningful; but for the case of Coulomb friction where singularities are reached after relatively short times, or static friction where the singularity is reached after a finite time span, the mathematical theory is insufficient. Nevertheless, for particles in contact, the linear oscillator with Coulomb friction is the prototype pattern of the flow in phase space.

1.5 Nonlinearities

Nonlinearities come up frequently in DEM simulations: even when linear interaction laws are assumed between contacting particles, the transition from non-contacting (zero force) to contacting (linear interaction) is nonlinear. The dynamics of nonlinear oscillators differ in various aspects from the dynamics of linear oscillators. The linear force law in Equation (1.92) contains no dependence of the period on the amplitude. For forces that are nonlinear and which grow more slowly than linearly in the displacement x, as in the case of the mathematical pendulum

$$m\ddot{x} + \sin(x) = 0, \tag{1.112}$$
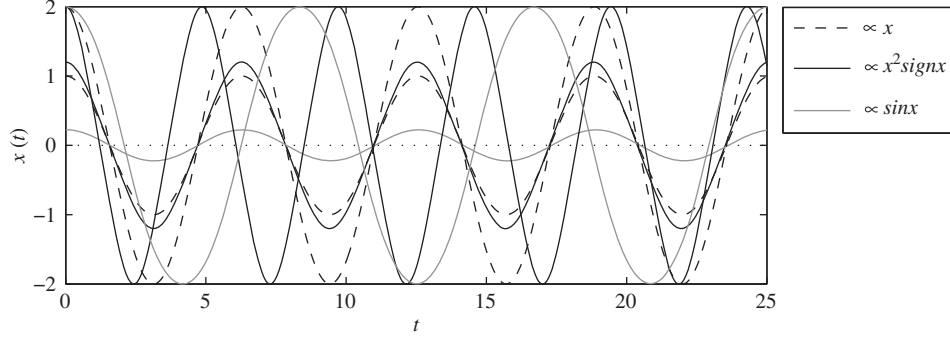
Figure 1.21 Relationship between period and amplitude for: the linear oscillator (dashed lines); the nonlinear oscillator of Equation (1.113) with $n = 2$ (solid black lines); and the mathematical pendulum of Equation (1.112) (solid gray lines). One curve of each pair has amplitude 2, and the other has amplitude such that the period is 2π.

the frequency decreases with the amplitude; see Figure 1.21. For forces that grow faster than linearly in the displacement x, such as

$$m\ddot{x} + |x|^n \, \mathrm{sgn}(x) = 0 \quad (n > 1), \tag{1.113}$$

the oscillation frequency increases with the amplitude; see Figure 1.21. The forces which result from particles coming into contact with deformations are not of this type: in that case, there is only a repulsive part, for wedge-shaped contacts with $n = 2$ and for spherically shaped contacts with $n = 3/2$; see Johnson [19]. Still, although the attractive part of the interaction is missing, the frequency dependence for half a period is important to know: when the amplitudes are high (e.g. high collision velocity, large compression), the frequency is higher and the time-scale is smaller; therefore smaller time-steps have to be used to resolve the corresponding particle contacts. In the same way as the frequency is influenced by the power of the displacement, the contact time for colliding DEM particles will be affected: in a temporary collision, instead of a full sine oscillation, only a single arc of the sine curve will be transversed by the contacting particles. Viewed in phase space, linear differential equations leave the shape of the cat's head as it is (Figures 1.18, 1.19 and the initial flow in Figure 1.20), whereas nonlinear differential equations distort the shape; see Figure 1.22 and the final stage of the flow in Figure 1.20.

1.5.1   Harmonic balance

The graphs in Figure 1.21 were produced via numerical integration. Analytical approaches to computing the amplitude dependence of the frequency are possible via the method of harmonic balance, i.e. by expanding the solution in a Fourier series (a sum of trigonometric functions) and considering the leading terms. For an oscillator with a third-order term (the Duffing oscillator)

$$m\ddot{x} + kx + \tilde{k}x^3 = 0,$$
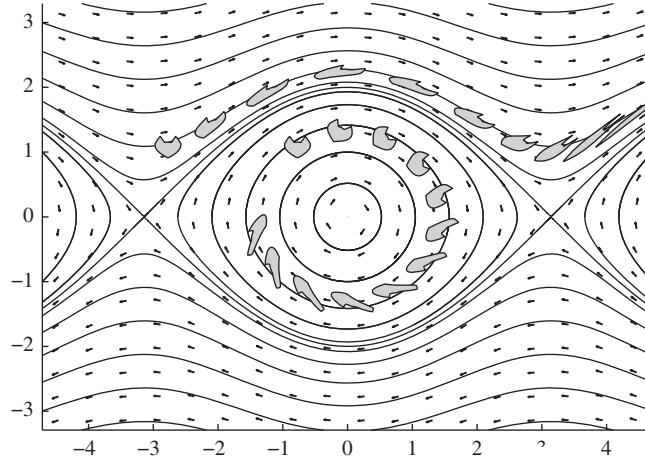
Figure 1.22 Phase flow for the mathematical pendulum of Equation (1.112): two phase space volumes chosen as cats' heads become distorted during their transport through phase space, due to the $\sin(x)$ nonlinearity in (1.112).

we first rewrite the equation as

$$\ddot{x} + \omega_0^2 x + \epsilon x^3 = 0, \tag{1.114}$$

and then approximate it by

$$\ddot{x} + \tilde{K}(x) = 0.$$

For small nonlinearities, we can assume solutions similar to those of the linear oscillator, in the form

$$\tilde{x}(t) = A \cos \omega t,$$

where instead of ω_0 we have to deal with the as-yet-unknown ω. Using the trigonometric identity

$$\cos 3\theta = 4 \cos^3 \theta - 3 \cos \theta,$$

we obtain the expansion of $\tilde{x}(t)$ in ω and its powers as

$$\tilde{K}(\tilde{x}(t)) = \omega_0^2 x + \epsilon x^3 = \left(\omega_0^2 + \frac{3}{4} \epsilon A^2 \right) A \cos \omega t + \frac{1}{4} \epsilon A^3 \cos 3\omega t.$$

Neglecting the third harmonic (the term with dependence on $3\omega t$) and substituting this expression for $\tilde{K}(\tilde{x}(t))$ into Equation (1.114) yields the linearization

$$\ddot{x} + \omega_0^2 \left(1 + \frac{3}{4} \frac{\epsilon A^2}{\omega_0^2} \right) x = 0. \tag{1.115}$$

So, for the nonlinear oscillator of Equation (1.114), the amplitude dependence of the frequency is approximately

$$\omega \approx \omega_0 \left(1 - \frac{3}{4} \frac{\epsilon A^2}{\omega_0^2} \right). \tag{1.116}$$

This agrees with the amplitude–frequency behavior in Figure 1.21 (though strictly speaking the nonlinearities in (1.113) with $n = 2$ and $n = 3/2$ cannot be expanded with leading terms in x^3): forces that grow faster than linearly lead to an increase in the frequency with increasing amplitude, while forces growing at a weaker rate lead to a decrease in the frequency with increasing amplitude. The resulting effect on the collision duration and the choice of time-step have already been discussed in the previous section.

1.5.2   Resonance in nonlinear systems

Resonance in nonlinear systems can be discussed analogously to the linear case in § 1.4.2. We make an ansatz for the solution,

$$x(t) = A \exp(i\Omega t), \tag{1.117}$$

where the frequency Ω is the frequency of the external excitation, and add damping so that from (1.114) we obtain

$$\ddot{x} + 2\delta\dot{x} + \omega_0^2 x + \epsilon x^3 = f_0 \exp(i\Omega t). \tag{1.118}$$

Here, due to the nonlinearity of the system, Ω will depend not only on the fundamental frequency ω_0, the nonlinear coefficient ϵ and the damping γ, but also on the amplitude A of the solution and the amplitude f_0 of the external excitation. Using the harmonic balance approach of the previous subsection, Equation (1.118) simplifies to
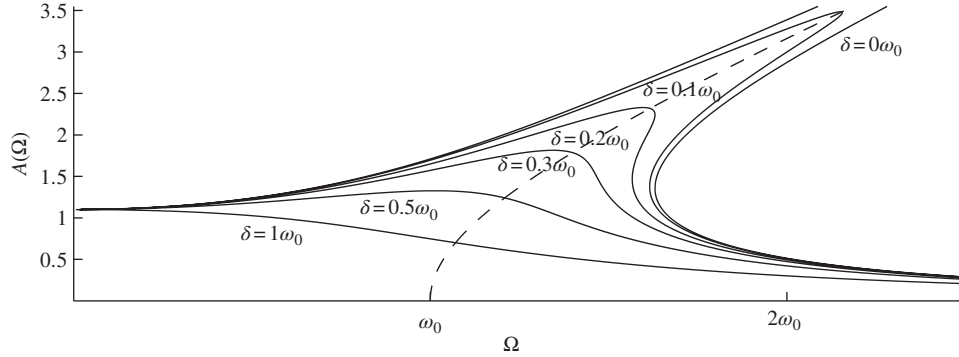
$$\ddot{x} + 2\delta\dot{x} + \omega_0^2 \left(1 + \frac{3}{4} \frac{\epsilon A^2}{\omega_0^2} \right) x = f_0 \exp(i\Omega t). \tag{1.119}$$

Plugging in the ansatz from (1.117), as in the linear case, we can eliminate the dependence on $\exp(i\Omega t)$ and get left with

$$-A\Omega^2 + 2\delta A\Omega i + \omega_0^2 \left(1 + \frac{3}{4} \frac{\epsilon A^2}{\omega_0^2} \right) A = f_0. \tag{1.120}$$

Instead of solving for A as a function of Ω, which would be a third-order equation, let us solve the second-order equation for Ω in terms of A. This gives the two solutions

$$\Omega_{1,2} = A\Omega^2 + 2\delta^2 - \frac{3}{4}\epsilon A^2 \pm \sqrt{\frac{f_0^2}{A^2} + 4\delta^2 \left(\delta^2 - \omega_0^2 + \frac{3}{4}\epsilon A^2 \right)}; \tag{1.121}$$

Figure 1.23 Curves of the relation between the amplitude A and the frequency Ω in harmonic balance for the parameters $\epsilon = 0.4\omega_0^2$, $f_0 = 1.5\omega_0^2$ and various values of δ in Equation (1.121). The dashed curve represents the relation between amplitude and frequency for free, undamped oscillations, i.e. the solution to (1.121) with $\gamma = 0$ and $f_0 = 0$. For smaller ϵ, the curve will be more upright; for negative ϵ, it will be tilted towards the left.
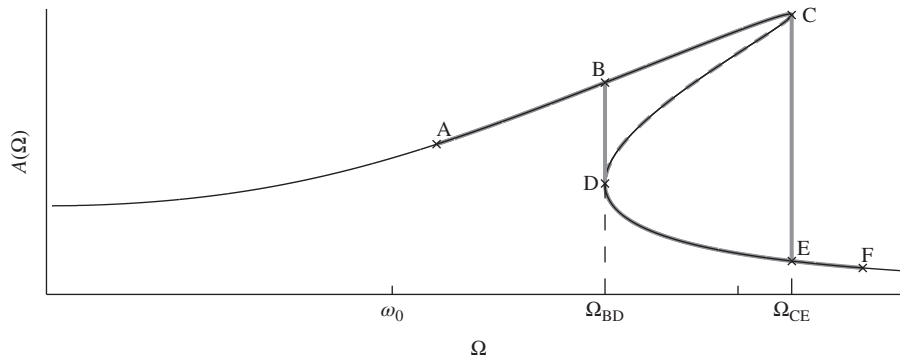
Figure 1.24 Hysteresis curve for the resonance of the nonlinear oscillator: when the frequency sweeps in a quasi-stationary way from lower to higher Ω, the resonance curve follows the path ABCEF, whereas from higher to lower Ω it follows the path FEDBA.

so the $\Omega_{1,2}$ indeed depend on all the other parameters in Equation (1.119). The graph for the real parts of $\Omega_{1,2}$ (only these are physically meaningful) is shown in Figure 1.23. Compared with the resonance curve for the linear oscillator, the cusp is tilted to the right for $\epsilon > 0$ (and it would be tilted to the left for $\epsilon < 0$). Because the solution for A in Equation (1.121) would be a third root, mathematically there are up to three solution points for a single value of Ω, i.e. there may be several amplitudes for the same frequency; which one of these is assumed by the system depends on the history. In Figure 1.24, possible transitions between the states are sketched: for an increase of Ω from point A to point F, the amplitude will follow the path ABCEF; for a decrease of Ω from point F to point A, the amplitude will follow path FEDBA. The amplitudes between B and D (gray dashed line in Figure 1.24) will usually not be assumed by the system. The phenomenon of different amplitudes being

selected in the range between Ω_{BD} and Ω_{CE}, depending on whether the control parameter Ω is increased or decreased, is called *hysteresis*. If we assume that a granular system is composed of particles with the nonlinear contacts described by the equations in this section, and if we assume that the whole system inherits the properties of the contacts, then for reasonably strong nonlinearities it becomes likely that certain vibration amplitudes cannot be realized: either too large or too small excitations take place. The authors have found such behavior in vibrated granular materials even in experiments: for some vibrated systems, convection was observed only for amplitudes that were larger or smaller than the actually desired amplitude, at which the system stood still.

1.5.3   Higher harmonics and frequency mixing

When we investigate physical systems, we input an external influence I (e.g. a force) and look at the response R of the system (e.g. the deformation); in the simplest case, there may be linear dependence

$$R = aI.$$

For a periodic input $I = \cos \omega t$ with frequency ω, the displacement will follow a temporal variation of the same frequency. When we have a nonlinear system, the nonlinear response can be expanded as a Taylor series; for example, to second order,

$$R = a_1 I + a_2 I^2. \tag{1.122}$$

The response to a periodic input $I = B \cos \omega t$ can then be rewritten via the trigonometric identity $\cos^2 \theta = \frac{1}{2}(\cos 2\theta + 1)$, with $\theta = \omega t$, as

$$R = a_1 B \cos \omega t + a_2 B^2 \cos^2 \omega t$$

$$= a_1 B \cos \omega t + \frac{a_2 B^2}{2} \cos 2\omega t + \frac{a_2 B^2}{2}. \tag{1.123}$$

In other words, the response will consist of a part with the original frequency ω, another part with doubled frequency 2ω, and a displacement from the original equilibrium $a_2 B^2/2$. A striking example from optics of second-harmonic generation by frequency doubling is the emission of blue light from an optically active target which is irradiated by a red laser of high intensity. For mechanical systems such as granular materials, we may also obtain an output spectrum that differs from the input spectrum (i.e. different frequencies, different wavelengths). There is another important consequence for disordered granular materials with nonlinear characteristics, i.e. particle contacts that obey nonlinear interaction laws: for sound waves with different finite amplitudes B passing through the same initial configuration, there may be a different reconfiguration of the granular matrix of magnitude $a_2 B^2/2$ in (1.123), which itself affects again the sound propagation; see [20] for a signature of such reordering in a DEM simulation of sound propagation through a system of poyhedral particles. Due to the strong frictional damping in granular materials, it will not be possible to reduce the amplitude too much, or else no output signal R can be measured at all. The generation of

higher harmonics, not only doubling the frequency, can be derived mathematically using the trigonometric identities

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta$$

for frequency tripling,

$$\cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1$$

for frequency quadrupling, and

$$\cos n\theta = n\cos^n\theta - \binom{n}{2}\sin^2\theta\cos^{n-2}\theta + \binom{n}{4}\sin^4\theta\cos^{n-4}\theta - \cdots$$
$$+ (-1)^k\binom{n}{2k}\sin^{2k}\theta\cos^{n-2k}\theta + \cdots$$

for the general case of nth-order harmonics. If there are oscillations with more than one input frequency ω, then there will be multiplicative terms; for example, with two input frequencies $A\cos\omega_1 t$ and $B\cos\omega_2 t$, the I^2 term in (1.122) will become

$$(A\cos\omega_1 t + B\cos\omega_2 t)^2 = A^2\cos^2\omega_1 t + 2AB\cos\omega_1 t\cos\omega_2 t + B^2\cos^2\omega_2 t,$$

so we have a product term $\cos\omega_1 t\cos\omega_2 t$. Using the trigonometric formula

$$\cos\theta\cos\phi = \frac{1}{2}\Big(\cos(\theta + \phi) + \cos(\theta - \phi)\Big),$$

we obtain 'sum frequency mixing' with $(\omega_1 + \omega_2)$ and 'difference frequency mixing' with $(\omega_1 - \omega_2)$. For many materials, nonlinear effects can often be 'argued away' based on small pre-factors. However, for the granular materials that we wish to study with the discrete element method, damping is often considerable, so one cannot work with small amplitudes, even in small laboratory experiments. Apart from that, some granular phenomena, such as landslides and earthquakes, naturally come with large amplitudes. Because of all these nonlinear effects which can modify the original frequency spectrum, it is not possible to rely on runtime experiments with mixed frequency spectra: waves $A(t, x, \omega)$ emitted at t_0 with a given frequency into the sample at one end may be damped out and not reach the detector at the other end of the sample at all; on the other hand, waves $B(t, x, \tilde{\omega})$ in between are generated from other frequencies and reach the target at times unrelated to t_0; see Shourbagy et al. [21] for a discussion of real data.

1.5.4   The van der Pol oscillator

An oscillator that exhibits some of the nonlinear frequency behavior discussed above is the forced van der Pol oscillator

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = \eta\sin(\omega t), \tag{1.124}$$
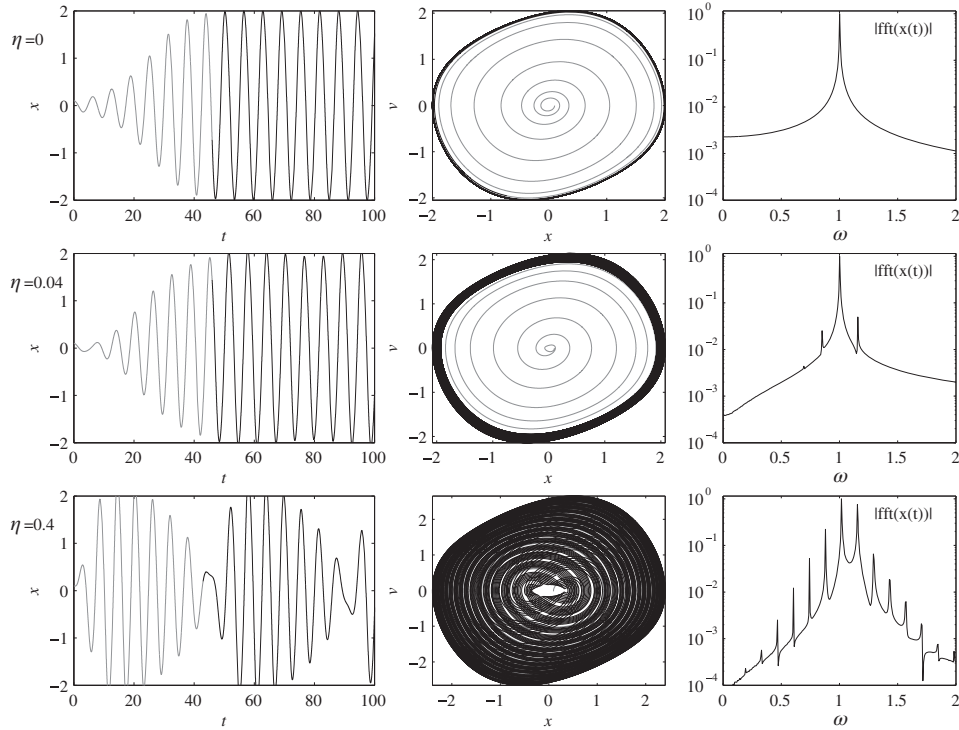
Figure 1.25   Position $x(t)$ (left column), phase portrait (middle column) and power spectrum with maximum value normalized to 1 (right column) for the van der Pol oscillator (1.124) with $\mu = 0.2$, $\omega = 1.15$ and the values of η shown at the left of each row. For computation of the power spectrum, the gray portions of the $x(t)$ curves (up to $t \approx 45$) were omitted, and data $x(t)$ with t up to about 1050 was used.

where the $x^2 \dot{x}$ term is nonlinear, of third order. The autonomous system (without explicit time dependence, where $\eta = 0$) oscillates with frequency $\omega = 1$ (see Figure 1.25, top row). The graph of $x(t)$ is not exactly sinusoidal, so the peak of the power spectrum, i.e. the absolute value of the Fourier transform (see § 5.2.2) is broadened around the fundamental frequency. When for $\mu = 0.2$ the external forcing is increased to $\eta = 0.04$, we see in the middle row of Figure 1.25 that another peak appears at a new, higher frequency, as well as a smaller one at $\omega = 0.85$ and a tiny one at $\omega \approx 0.7$. This indicates the presence of difference frequency mixing. For larger forcing with $\eta = 0.4$, the difference mixing spreads out over the whole spectrum, and the peak at $\omega = 1$ nearly reaches the amplitude for the eigenfrequency $\omega = 1$ of the unforced oscillator (Figure 1.25, bottom row). The Poincaré–Bendixson theorem prohibits the occurrence of chaos (in the exact mathematical sense) in a continuous dynamical system in the plane, so the van der Pol oscillator (which has only two coordinates, x and v) can have only a discrete spectrum. The Fourier transform used in the right column of Figure 1.25 gives a peak in the spectrum for ω. The peak is of finite width, even for the unique stable

trajectory which exists when $\eta = 0$. Therefore, the Fourier transform is not the optimal tool for analyzing whether a spectrum is continuous or not. In the next section, we will discuss a model which indeed exhibits a continuum of states, and introduce an method of analysis that does not require use of the Fourier transform.

1.6   From higher harmonics to chaos

1.6.1   The bifurcation cascade

As the strength of the nonlinearity increases, phenomena can occur which are unexpected from the point of view of 'linearized' mechanics. Such phenomena can affect the observable computation and the accuracy with which specifications for experiment and simulation have to be given, and there may be considerable scattering of data even when the initial conditions are 'nearly identical' or if the system is perturbed a little to have 'slightly different' dynamics. Even if only the time-step of computation is changed, for large enough particle numbers the configuration may evolve along totally different trajectories. While there are many treatises on chaos in mechanics, few are directly applicable to DEM simulations; here we shall give an overview of phenomena that can actually affect the development of DEM programs. The generation of higher harmonics means that for a given discrete spectrum of input frequencies $S_{\text{in}} = \{\omega_1, \omega_2, \dots\}$, the system could respond with an output spectrum $S_{\text{out}} = \{\tilde{\omega}_1, \tilde{\omega}_2, \dots\}$ that is different but still discrete. Beyond that, there is a possibility of going from discrete to continuous spectra in a bifurcation scenario: as the nonlinearity parameter (called η in the following) increases, the response parameter could split into two branches repeatedly, until a continuum of states ('chaos') is reached; see Figure 1.26.

1.6.2   The nonlinear frictional oscillator and Poincaré maps

We consider the differential equation for the nonlinear friction oscillator ('stick-slip oscillator'),

$$\ddot{x} + x + a\left[\mu(1) + \mu(\dot{x} - 1)\,\mathfrak{sgn}\,(\dot{x} - 1)\right] = \gamma\cos(\eta t), \tag{1.125}$$
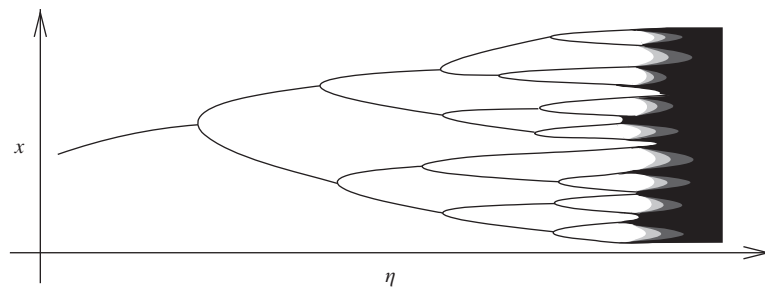
Figure 1.26   Bifurcation scenario of a variable x, showing successive (not necessarily symmetric) period-doubling up to a continuum of states as an external nonlinearity parameter η is increased, leading to the development of chaos.
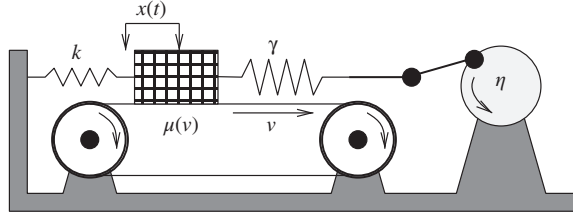
Figure 1.27 A physical system with the behavior of the differential equation (1.125): a mass on a conveyor belt with several couplings, which alternates between sliding and sticking.
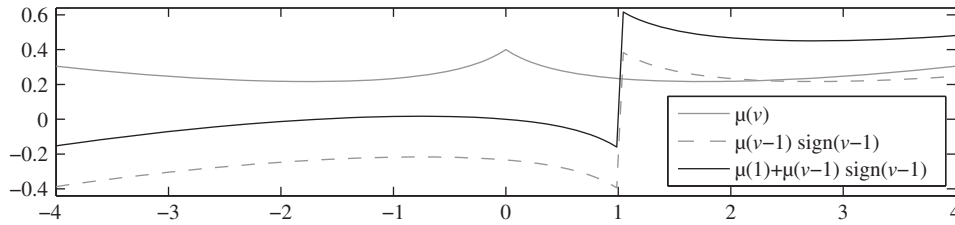
Figure 1.28 Graphs of the velocity-dependent friction law $\mu(v)$ of (1.126) and different combinations of terms in the Stribeck friction expression in Equation (1.125).

where we use η rather than ω to represent the frequency, indicating that it will be our nonlinearity parameter. Here we take the sign function of (1.98), $a = 10$, and a velocity-dependent friction law with

$$\mu(v) = \frac{\mu_0 - \mu_1}{1 + \lambda_0|v|} + \mu_1 + \lambda_1|v|^2, \tag{1.126}$$

where the pre-factors of the velocity dependence are $\lambda_0 = 1.42$ and $\lambda_1 = 0.01$. The coefficient of static friction, $\mu_0 = 0.4$, is larger than the coefficient for dynamic friction, $\mu_1 = 0.1$. The velocity dependence is sketched in Figure 1.28: note that $\mu(v)$ is symmetric in v; the physical dependence of the sign requires the multiplication with $\mathfrak{sgn}(\dot{x} - 1)$ in Equation (1.125). Velocity-dependent characteristics similar to $\mu(v)\,\mathfrak{sgn}(x)$ are sometimes referred to as 'Stribeck friction'; an example is the friction between violin strings and the rosin-coated violin bow [22, p. 284]. A physical system corresponding to Equation (1.125) is depicted in Figure 1.27: a mass connected to a spring with spring constant 1 slips or sticks on a belt, with the mutual friction given by (1.126).

The frictional oscillator of Equation (1.126), like the van der Pol oscillator, has only a single position coordinate x and a single velocity coordinate v, but it is not a purely two-dimensional system. Our at-first-glance elusive definition of the sign function in (1.98), which leaves λ in the range $-1 \leq \lambda \leq 1$ such that the external force can be compensated, in fact includes an additional parameter (the 'Conley index'; see Kunze [23]) which can act as a further dimension to the problem. In §3.3.2, we will show how the computation of λ can be performed in a 'numerically exact' manner (i.e. with controllable discretization errors,

and without any modeling assumptions). Because of the variation of λ, the system exhibits true chaos without running afoul of the Poincaré–Bendixson theorem, which forbids chaos in purely two-dimensional continuous systems. For discrete systems, even in one dimension, such as for the logistic map

$$x_{n+1} = \eta x_n (1 - x_n),$$

chaotic behavior (in the sense of continuous distributions for the x_{n+1}) is possible; see § 1.6.3. The corresponding continuum model, the logistic equation

$$\frac{dx}{dt} = \eta x (1 - x),$$

has the explicit solution

$$x(t) = \frac{1}{1 + \left(\frac{1}{x_0} - 1\right) \exp(\eta t)},$$

which is not chaotic at all. This should serve as a warning to anyone who tries to model the physical behavior of systems of discrete particles with continuum approaches: the same dynamics is not necessarily accessible when one goes from discrete to continuous models in a given dimension. The solutions to the nonlinear friction equation (1.125) vary strongly with η, as can be seen from the equilibrium trajectories in Figure 1.29 (i.e. trajectories omitting the initial part of the solution); depending on η, the solutions may differ considerably.

One might guess that the solution is periodic, or not. To make it easier to investigate the periodic dynamics of the system and its dependence on the parameter η, instead of looking at the Fourier transform as in § 1.5.4, we will investigate the Poincaré map (or Poincaré cut, as it is obtained as an intersection with the plane at a given η), which is the intersection of the trajectory in phase space with a plane defined at a certain velocity (Poincaré section); see Figure 1.30. This reduces the effective dimension of the system by 1. Instead of n peaks in the Fourier transform, for suitable chosen (half-)planes (we will choose $v = 0$
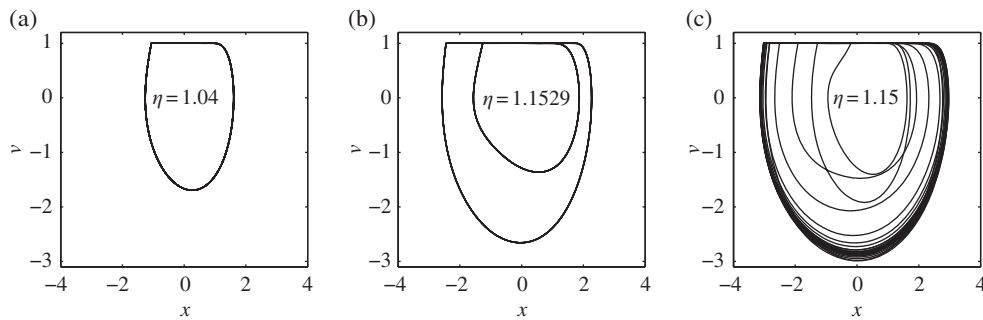
Figure 1.29 Some trajectories for different values of the parameter η in the nonlinear friction oscillator equation (1.125). The cats' heads are not shown, but they would all be contracted on the lines for the trajectories.
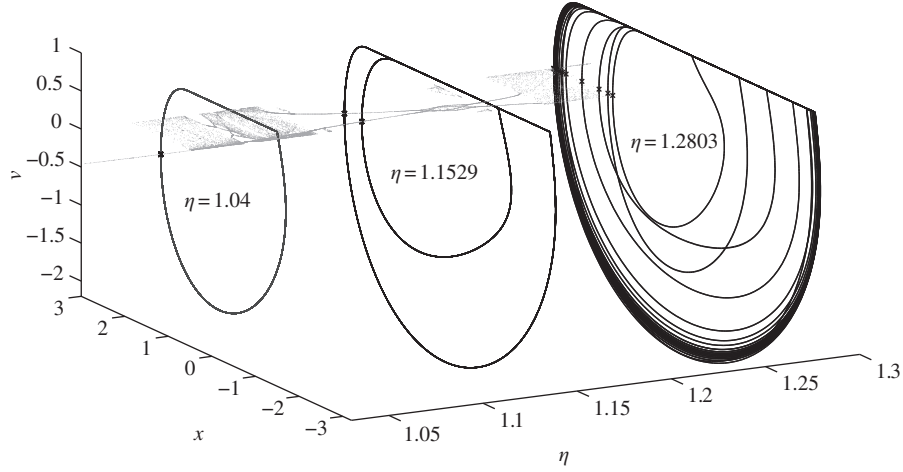
Figure 1.30 Selected trajectories for various values of the parameter η in the nonlinear friction oscillator equation (1.125). Intersection points between the trajectory $x(t)$ and the plane $v = 0$ are marked by crosses; further (numerically computed) intersection points for this Poincaré map are marked by gray dots and are replotted in Figure 1.31 in two dimensions.
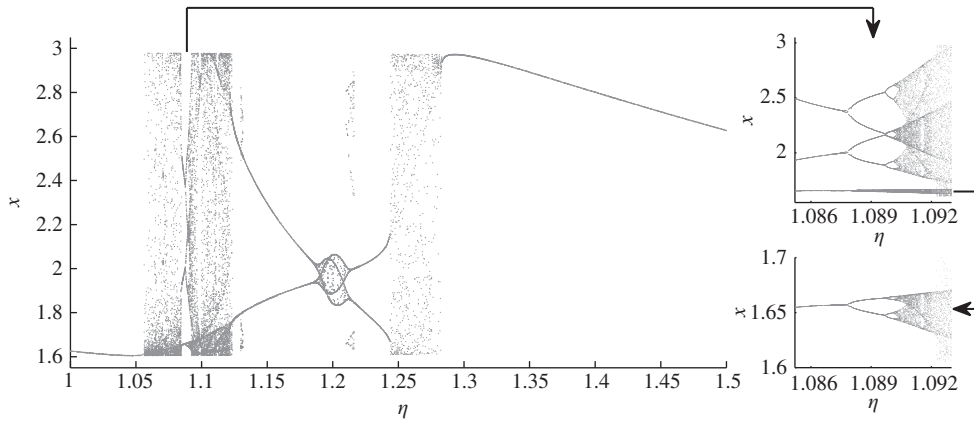
Figure 1.31 Return map (Poincaré map) for the nonlinear friction oscillator of (1.125), obtained from the Poincaré section at $v = 0$. Values of $x(t) \geq 0$ are plotted for different values of η; the two insets display successively magnified phase space volumes to show the fine structure.

for $x > 0$) one finds n intersection points between the trajectory and the plane. The return map is plotted in Figure 1.31: the nonlinear frictional oscillator alternates between oscillating among a set of discrete values (periodic dynamics) and visiting a range of practically continuously distributed values (chaos) at different values of η. This alternating behavior is called intermittency. The practically continuous spectrum is a sign of mechanical chaos: initially close trajectories can diverge arbitrarily far. Attractors exhibiting this kind of behavior

are also called 'strange attractors'. Chaos is the most highly nonlinear form of nonlinearity. The short-term behavior is predictable, but the long-term behavior is not. Although trajectories are unpredictable, there is a definite mathematical structure that allows one to predict in which parameter region chaos will occur. Despite this, due to finite errors which are inherent in modeling a system, one may obtain practically random behavior from systems which are to all intents and purposes deterministic. Note, however, that although the distribution of values in a chaotic system is continuous, it is by no means uniform, as can be seen from the shading in Figure 1.31. Because, moreover, the order in which the continuum is sampled is difficult to conceive, Poincaré maps of the chaotic regime cannot be used as, for instance, random number generators, for which there are better alternatives (see, e.g., vol. 2 of [24]).

1.6.3   The route to chaos

The sequential growth in complexity of the dynamics with the strength of the nonlinearity, as in the bifurcation scenario, is sometimes called the 'route to chaos'. For a linear system, there is a single mode (e.g. velocity, frequency, wavenumber, position, or a combination of these). When nonlinearity is involved, additional peaks can be observed in the spectrum. Eventually there is a transition from a spectrum of densely positioned peaks to a continuous spectrum—to chaos. Even in the chaotic case it does not mean that the probability density of each trajectory is the same, as can be seen from Figure 1.31, which definitely shows structure even in the chaotic region. The classical bifurcation scenario assumes that each stage involves a doubling of peaks, but that is not what we see for our nonlinear frictional oscillator in Figure 1.31. Some return maps are self-similar, or fractal; this is the case for the final (stationary) values of the discrete iteration known as the 'logistic map':

$$x_{n+1} = \eta x_n(1 - x_n) \tag{1.127}$$

with nonlinearity parameter η. Self-similarity means that if one magnifies a portion of the diagram, one sees basically (and in some cases, after transformation of the axes, exactly) the same overall structure as the original; see Figure 1.32, where successively magnified portions of the map are shown. The Poincaré map of the frictional oscillator in Figure 1.31 is not fractal; the Coulomb friction seems to break the scale-invariance that is inherent to the return maps of many nonlinear systems. This means that one has to be careful when adapting concepts of nonlinear theory to realistic mechanical systems, especially granular materials. Particle size, friction and other physical properties lead to characteristic dynamics at different scales, which may be incompatible with aspects of nonlinear systems such as self-similar return maps.

While chaos itself inhibits the computation of individual trajectories in accordance with experimental data, it may actually be an asset for the theoretical prediction of statistical properties of many-particle systems. Molecular chaos, the assumption that velocities of colliding particles are uncorrelated and independent of position (Boltzmann's 'Stosszahlansatz', or collision-parameter approach) underlies many analytical methods for collision-dominated particle systems, including granular particles at low densities. In fluid mechanics, chaos is equivalent to turbulence, i.e. a continuous size distribution of vortices from the largest to the smallest length scales. In fluid dynamics, 'routes to chaos' via bifurcation can evolve simultaneously in the same system at different places: for the separation flow in a transitional boundary layer with an impinging shock-wave as external forcing, the spatial and temporal
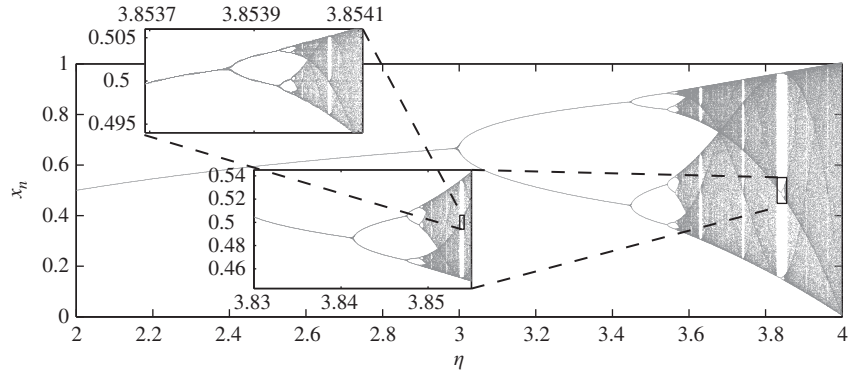
Figure 1.32 Return map for the logistic map of Equation (1.127). Successive inserts show magnified detail of the previous map; while having different scales, all three plots show the same structure, demonstrating the fractal (self-similar) nature of the map.

development of the first vortex at the impinging point towards the vortex field further downstream (see [25]) follows the bifurcation cascade in Figure 1.26. It cannot be excluded that different stages of the development of chaos might occur simultaneously in granular systems simulated with the discrete element method.

1.6.4   Boundary conditions and many-particle systems

The character of the nonlinearity may be not only a matter of the dynamics of the physical system but also of the boundary conditions. In Figure 1.33 we contrast the trajectories in a conventional billiard geometry and in a 'stadium billiard' geometry, for constant absolute velocity; the dynamics is that of a single particle which gets reflected at the boundaries. In the conventional billiard case, the trajectories are parallel, whereas in the stadium billiard case they diverge and, for certain types of boundaries, become chaotic [26]. This means that sharpness of corners (here, of the system boundaries) is not in itself a guarantee of the existence of nonlinearities. As the divergence of initially close trajectories may be desired, for example when considering mixing in hoppers, one has to pay proper attention to the shape of the boundaries. For sharp corners rather than flat edges, the character of the nonlinearity can be assumed to increase. Especially for particle systems with low density, boundary and initial conditions will have considerable influence on the dynamics, beyond mere interaction. For the simulation of accretion disks via smoothed particle hydrodynamics (SPH), a symmetric choice of initial positions and velocities has been found to cause axisymmetric stripes at a later stage in the simulations [27], which overlay the inherent instabilities of the system [28]. As SPH uses more interaction partners and stronger averaging than the discrete element method, one has to be even more careful with 'harder' (more nonlinear) interactions than in DEM simulations.

    Chaos can easily occur in mechanical multi-body systems: already the double-pendulum, which has only two degrees of freedom, can exhibit chaos [29]. In dry granular materials, there are several aspects which contribute to the nonlinear character on the level of individual particles: the first is the transition from no interaction for separated particles to a finite interaction
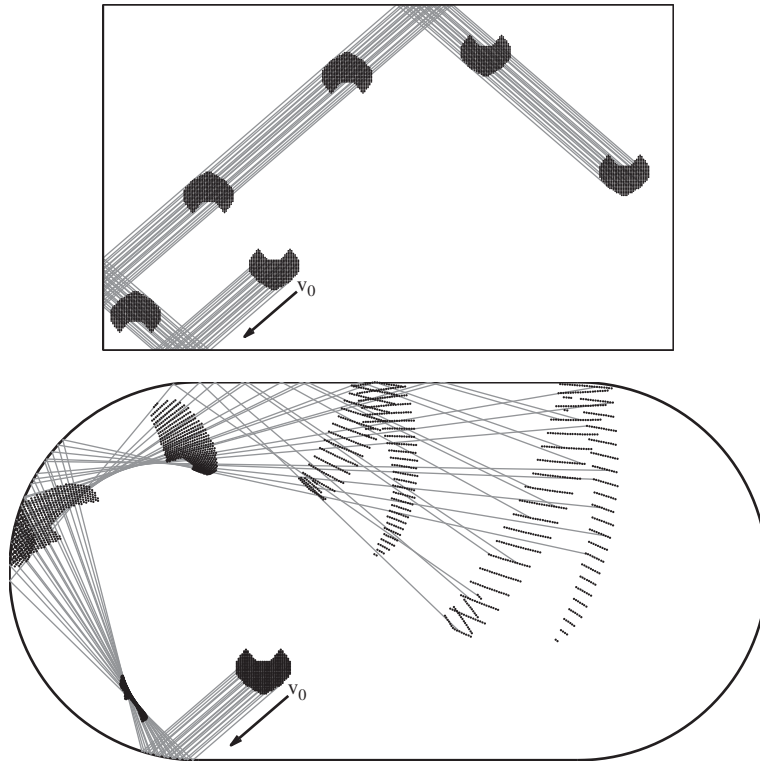
Figure 1.33 Effect of the boundary condition on the nonlinearity of the system (manifested here as the divergence of trajectories): trajectories of a system with rigid reflection when the boundary conditions are shaped according to a conventional billiard table (above) or a 'stadium billiard' table (below). The same initial velocity is assumed in both cases; the time evolution of points from the set of initial conditions (cat's head) is shown in black, while selected trajectories are drawn in gray.

for particles in contact, which may actually be more decisive than the second aspect, which is the detailed nonlinear power of the interaction. Computationally, chaos was discovered by E. Lorenz, who found, in a nonlinear oscillator system with three variables, wildly different solution trajectories from only slightly different initial conditions [30]. One should not be surprised to encounter this behavior in discrete element solutions as well. The Euler equations of motion for rotation, (1.35)–(1.37), are themselves nonlinear; we have observed in polyhedral particle simulations that minimal changes (even a mere reduction in the step-size or choice of a different order in the summation of the forces) could lead to a strong divergence of the orientation of the particles, although at least in the beginning the positions of the center of mass were not affected.

1.7 Stability and conservation laws

Stability is the notion that a system 'does not change much' under a perturbation. This means that if we repeat the same experiment (or calculation, or simulation) with slightly different initial conditions, the outcome should also not change much. Here we review some basic ideas
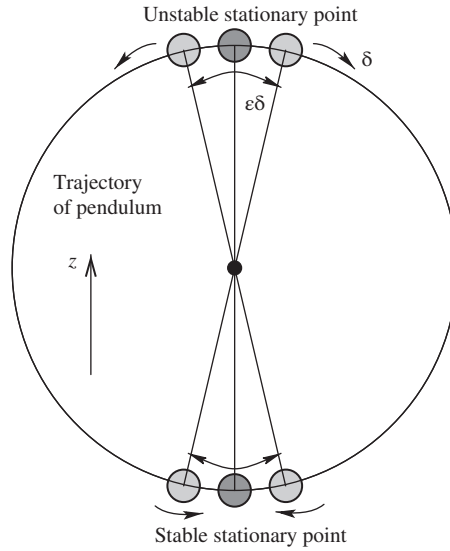
Figure 1.34   Stability and instability for a pendulum resting at different stationary points. After a small displacement δ from the upper stationary point, the bob moves away from the position, which is therefore an unstable state. The bob always returns to the lower stationary point after a small displacement, so this position is a stable equilibrium.

from stability theory, but will not go into details: while the general notion is important for simulations of mechanical systems, almost all DEM systems will turn out to be unstable in the sense of classical stability in mechanics, which was devised more with celestial mechanics in mind than with the aim of describing friction- and dissipation-influenced phenomena on Earth. Nevertheless, the concept (though not the mathematical theory) of stable and unstable quantities is useful in helping us focus on appropriate observables in particle simulations. Further, we outline which conservation laws are suitable for testing the quality of DEM simulations.

1.7.1   Stability in statics

Mechanical stability, or lack thereof, is usually defined with respect to (not necessarily one-dimensional) stationary points x_s (also called equilibrium positions [31, p. 797]) of a physical system. If for $x(t) = x_s$, $v(t) = 0$, the position will stay at x_s always, so x_s is said to be a stationary point. If after a small deflection δ from x_s, the system stays close to x_s, then the stationary point is stable; if the system moves away after a small deflection δ, the stationary point is unstable [32, p. 166]. The bob of the pendulum in Figure 1.34 has two stationary positions, at the highest and lowest points; the upper one is unstable and the lower one is stable.

The formal definition of stability is that for any $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that whenever

$$|x(0) - x_s| < \delta, \tag{1.128}$$

we have

$$|x(t) - x_\mathrm{s}| < \epsilon \tag{1.129}$$

for all $t > 0$. This means that solutions which start 'close enough' to the equilibrium (within a distance δ of it) remain 'close' forever (within a distance ϵ). Note that this must be true for any infinitesimal $\epsilon > 0$ that one might choose. Equations (1.128) and (1.129) have the same mathematical form as the Weierstrass criterion for continuous functions (epsilon-delta continuity) [31, p. 57]: a function $f(t)$ is continuous if for any $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that whenever $|t - a| < \delta$, we have

$$|f(t) - f(a)| < \epsilon. \tag{1.130}$$

1.7.2 Stability in dynamics

We define stability in dynamics analogously to the stability of points in statics, by generalizing from single positions to entire time-dependent trajectories $x(t)$, i.e. to solutions of an initial value problem for a differential equation. Let t_i denote the initial time and t_f the final time of interest. We write $q(t_\mathrm{i}) = q_\mathrm{i}$ and $\tilde{q}(t_\mathrm{i}) = \tilde{q}_\mathrm{i}$ for two initial states, and $q(t_\mathrm{f}) = q_\mathrm{f}$ and $\tilde{q}(t_\mathrm{f}) = \tilde{q}_\mathrm{f}$ for the final states on the corresponding trajectories. Then we have stability if for initially close coordinates with

$$|q_\mathrm{i} - \tilde{q}_\mathrm{i}| < \delta, \tag{1.131}$$

the separation between the final coordinates is bounded by a function which is a power of the time span:

$$|q_\mathrm{f} - \tilde{q}_\mathrm{f}| < \epsilon \, (t_\mathrm{f} - t_\mathrm{i})^p. \tag{1.132}$$

If the deviation diverges exponentially,

$$|q(t) - \tilde{q}(t)| = C \exp(\lambda t) \tag{1.133}$$

for some $\lambda > 0$, the system is said to be Lyapunov unstable, and λ is called the Lyapunov exponent; see Figure 1.35. From this follows a definition of stability via 'Lyapunov functions': if a solution can be constructed using exponential functions with positive exponents, then it is unstable. Because of the symmetry between coordinates and their velocities, as mentioned in § 1.4, velocities may be included in the norm (i.e. distance) measurements of (1.131)–(1.133).

Strictly speaking, these definitions are valid only for trajectories which correspond to solutions of ordinary differential equations that are autonomous systems without dissipation. We have seen in the previous discussions on resonance that a periodic perturbation can generate infinite amplitudes in the absence of damping, which makes the definition of stability more complicated; see Merkin [33, p. 226ff]. Furthermore, one should really make a distinction between a theory for finite times and one for infinite times (see the Introduction of [33]). Earlier we introduced some other mathematical phenomena that can lead to instability; for example, chaos, whereby the Poincaré map generates a continuum of points, is a type of

(a)

(b)
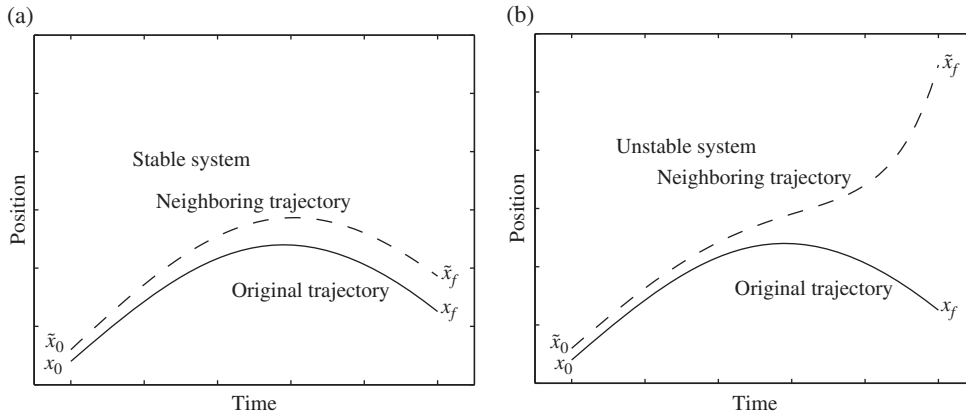
Figure 1.35   Neighboring trajectories of: (a) a stable system; (b) an unstable system.

instability. Certain kinds of attractors are also indicators of instability; see the elaborate discussion in Greiner [34, p. 467ff]. If dissipation is added to stable mechanical systems, the stability of otherwise stable structures can be destroyed [33, p. 202], which is rather counterintuitive: one might expect that dissipation, which removes energy from the system and reduces particle motion, would increase stability. However, as demonstrated by the hysteresis jumps in the resonance curve with linear damping of Figure 1.24, damping can indeed lead to a loss of stability. On the other hand, that the usual definitions of stability result in dissipative systems being classified as generally unstable reflects the fact that these ideas of stability originate from celestial mechanics and have only limited applicability to terrestrial mechanics. We might hope for a definition that could discriminate between 'stable' and 'unstable' slopes in granular materials; however, no such theory exists.

There are other aspects of classical stability theory that make its application to particle mechanics problems, such as the discrete element method, difficult. For instance, displacements (or the corresponding perturbations to the systems) are always 'infinitesimal' in mathematical stability theory; but for real systems and finite displacements, doubt remains as to whether this mathematical theory can describe appropriately the actual stability or instability of mechanical systems. Arnold [13, p. 121] gave a mathematical proof that while the acrobatic (inverted) pendulum (i.e. the bob positioned at the apex of the trajectory in Figure 1.34) is unstable, the same stationary point becomes stable if the pendulum is vibrated vertically. An experimental realization of this scenario would be balancing a pencil by merely moving it up and down in one's hand. This would certainly not be sufficient to keep the pencil upright, so in this case mathematical rigor is not the same as physical relevance. As any perturbations in physical experiments are finite, but the mathematical theory assumes infinitesimal perturbations, we would still consider the system in question as being unstable based on everyday experience.

1.7.3   Stable axes of rotation around the principal axis

The previous discussions pertained to rectilinear degrees of freedom. For angular motion, where different degrees of freedom are coupled, a slightly different approach than the $\epsilon - \delta$

philosophy is needed. We show here the common analysis of which axes are stable for rotation around the principal axis, i.e. the axis obtained from the coordinate transformation in such a way that the tensor of inertia is diagonal. For the torque-free case, we can rewrite the Euler equations of motion, (1.35)–(1.37), in terms of the angular momenta L_i^b. To do this, multiply the first Euler equation by $J_2 J_3$, the second by $J_1 J_3$ and the third by $J_1 J_2$ (where J_1, J_2, J_3 are the diagonal elements of the tensor of inertia), then replace each $J_i \omega_i^b$ with L_i^b to obtain

$$J_2 J_3 \dot{L}_1^b = (J_2 - J_3) L_2^b L_3^b, \tag{1.134}$$

$$J_1 J_3 \dot{L}_2^b = (J_3 - J_1) L_1^b L_3^b, \tag{1.135}$$

$$J_1 J_2 \dot{L}_3^b = (J_1 - J_2) L_1^b L_2^b. \tag{1.136}$$

Now, if we multiply (1.134) by $J_1 L_1^b$, (1.135) by $J_2 L_2^b$ (1.136) by $J_3 L_3^b$, and add up the resulting equations, we obtain

$$J_1 J_2 J_3 \left(L_1^b \dot{L}_1^b + L_2^b \dot{L}_2^b + L_3^b \dot{L}_3^b \right) = 0$$

Integration with respect to time then gives

$$\left(\left(L_1^b \right)^2 + \left(L_2^b \right)^2 + \left(L_3^b \right)^2 \right) = C_1, \tag{1.137}$$

where C_1 is a constant, equal to the value of L^2, the square of the total angular momentum L^b of the object in the body-fixed coordinate system. So we have shown that in the absence of an external torque, the total angular momentum is conserved.

If we multiply (1.134) by L_1^b, (1.135) by L_2^b and (1.136) by L_3^b and then add up the resulting equations, we obtain

$$\left(J_2 J_3 \left(L_1^b \right)^2 + J_1 J_3 \left(L_2^b \right)^2 + J_1 J_2 \left(L_3^b \right)^2 \right) = 0.$$

Integrating with respect to time and dividing by $J_1 J_2 J_3$ gives

$$\left(\frac{\left(L_1^b \right)^2}{J_1} + \frac{\left(L_2^b \right)^2}{J_2} + \frac{\left(L_3^b \right)^2}{J_3} \right) = C_2. \tag{1.138}$$

Recall that in the context of rectilinear degrees of freedom, for mass m, momentum p and velocity v we have $p^2 = m^2 v^2 = 2T$, where T is the kinetic energy; so we see that C_2 must correspond to twice the kinetic energy of the intrinsic angular momentum T^i, which is also conserved. At the same time, we see that for the individual components 1, 2, 3 of kinetic energy and angular momentum, no conservation law can be derived, as the Euler equations of motion couple the three components together. Equation (1.138) has the same functional form as the general equation for an ellipsoid,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$
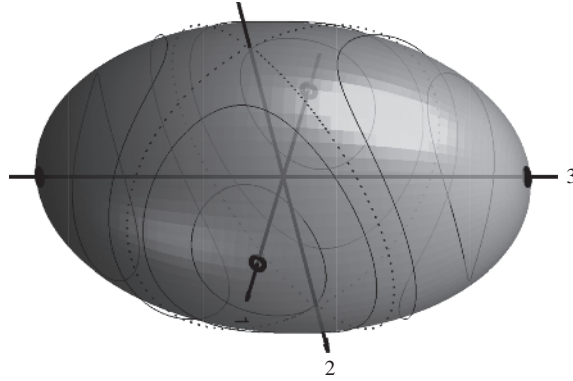
Figure 1.36   Poinsot ellipsoid of constant kinetic energy for the intrinsic angular momentum; stable trajectories around the principal axes 1 and 3 are plotted as solid lines, and the trajectory around the unstable principal axis 2 is plotted as a dashed line.

for which the half-axes of lengths a, b and c are aligned parallel to the Cartesian axes x, y and z, respectively. The ellipsoid described by Equation (1.138) is called the Poinsot ellipsoid, after L. Poinsot, who first proposed an interpretation of rotation as rolling of this ellipsoid on a plane in angular momentum space [35]. Let us order the axes so that $J_1 < J_2 < J_3$. For fixed L^b, the kinetic energy for rotation around axis 1 is $T^i_{\min} = L^b/(2J_1)$, which is maximal, while the kinetic energy for rotation around axis 3, $T^i_{\max} = L^b/(2J_3)$, is minimal. Accordingly, the kinetic energy T^i is bounded by

$$T^i_{\min} \leq T^i \leq T^i_{\max}.$$

In Figure 1.36 we plot trajectories of **L** for various initial conditions, along with periodic 'trajectories' of the angular momentum. It turns out that rotations around axis 1 (corresponding to the smallest moment of inertia J_1) and around axis 3 (the largest moment of inertia J_3) are stable, whereas rotations around axis 2 (intermediate moment of inertia J_2, where $J_1 < J_2 < J_3$) are unstable. It is also possible to prove the Lyapunov instability analytically; see [36]. The assumption of energy conservation is rather strong—too strong to be valid for many technical applications. If a system is energy dissipating, a reduction in the energy will force the rotation to be around axis 3 with the minimum kinetic energy and maximum moment of inertia. Even satellites have been found to show enough dissipation so that axis 1 with the minimal moment of inertia and therefore the maximal kinetic energy becomes unstable [37, p. 62ff]. This shows that stability proofs for 'ideal' (e.g. frictionless) systems are not all that relevant for technical systems, including particle systems and the discrete element method.

1.7.4   Noether's theorem and conservation laws

Noether's theorem (see, e.g., [38, p. 359]) asserts that where there is a symmetry in a mechanical system, there is a conserved quantity. From the homogeneity of time (every time interval looks like every other time interval) follows the conservation of total energy. From the homogeneity of space (every spatial interval looks like every other spatial interval) follows the

conservation of momentum. From the isotropy of space (every direction looks like every other direction) follows the conservation of angular momentum. Such conservation laws can be used to test, for example, whether interaction laws have been implemented correctly in models. It will be useful to discuss briefly what quantities can be conserved for discrete element systems. (It is advisable to test such conservation laws first with two particles in the system to avoid losing the overall view of the simultaneous interactions.) The total energy in a particle system consists of kinetic energy, both for the rectilinear degree of freedom, (1.26), and for the angular degree of freedom, (1.27), and potential energy. In the body-fixed system, the position energy for particle systems consists of external potentials, usually gravity, and interaction energy, which for DEM particles would be 'elastic' energy due to overlapping or deformed contacts. While for gravity the potential energy is easy to compute, in the DEM case potential energy can be difficult to estimate for anything other than linear potentials. Forces F that are 'conservative', i.e. which conserve the total energy of a particle moving under their influence) are those which can be derived from the gradient of a scalar field Φ, so that

$$F = -\nabla\Phi.$$

Besides for gradient potentials, energy conservation holds for rotationally symmetric potentials; for collisions with other potentials, a violation of energy conservation must be expected. In other words, for many discrete element models with non-spherical particles, there is no energy conservation even in the absence of velocity-dependent forces. Imagine the one-dimensional propagation of a particle through a force field which is asymmetric: from x_1 to 0 the force increases linearly, and from 0 to x_2 it decreases linearly, so that it is described by the formula

$$F(x) = \begin{cases} 0 & \text{for } x < x_1, \\ k_1(x - x_1) & \text{for } x_1 < x < 0, \\ k_2 x - k_1 x_1 & \text{for } 0 < x < x_2, \\ 0 & \text{for } x > x_2, \end{cases} \tag{1.139}$$

where $|k_1| \neq |k_2|$, $k_1 > 0$, $k_2 < 0$ and $|k_1 x_1| = |k_2 x_2|$; see Figure 1.37. Consider a particle moving from a position $x_3 < x_1$ to another position $x_4 > x_2$. Because the work performed on the particle is

$$W = \int_{x_3}^{x_4} F(x)\, dx > 0, \tag{1.140}$$
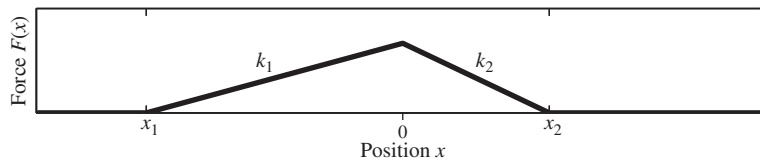
Figure 1.37 Sketch of the asymmetric force field given by Equation (1.139), where k_1 and k_2 are the gradients of the force at different positions.

Space

Time
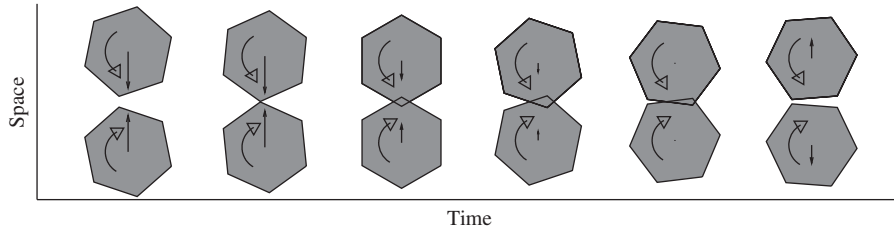
Figure 1.38 One-dimensional collision of non-spherical particles: the kinetic energy of the rectilinear degree of freedom is greater before than after the collision, even if there are no velocity-dependent damping forces. At the approach of the contact, the interaction is via a wedge–wedge contact (proportional to d^2 where d is the deformation; see § 1.5); on separation the interaction is linear (proportional to d^1).

the kinetic energy of the particle at x_4 will be different than at x_3. At x_3 and x_4, the kinetic energy is the same as the total energy, because there is no other potential. So particles which traverse asymmetric force fields on approach may undergo a change of their total energy even in the absence of velocity-dependent dissipative forces. This often happens in particle simulations of collisions involving rotations of non-spherical particles: when the particles turn during the collision, the repulsive force on approach can be different from the force upon separation; see Figure 1.38. This is one reason why verifying conservation of energy is not a useful way to check whether the simulation and time integrator were implemented correctly; another reason is the time integration method itself (see Chapter 2, in particular § 2.4). Only for spherical particles, owing to the rotational symmetry, is the energy conserved with certainty.

Nevertheless, for asymmetric forces (without velocity-dependent dissipative forces), time reversal is a suitable test. First, one runs the particle collision forward from initial time t_i to final time t_f:

$$(\vec{x}_i, \mathbf{q}_i, \vec{v}_i, \boldsymbol{\omega}_i) \rightarrow (\vec{x}_f, \mathbf{q}_f, \vec{v}_f, \boldsymbol{\omega}_f);$$

then one runs the process backward and compares the respective increase and loss of energy, which should add up to zero. In practical terms this means running the integrator backward in time (with negative time-step) or running the integrator with positive time-step but with the final velocities reversed, i.e.

$$(\vec{x}_f, \mathbf{q}_f, -\vec{v}_i, -\boldsymbol{\omega}_f) \rightarrow (\vec{x}_i, \mathbf{q}_i, -\vec{v}_i, -\boldsymbol{\omega}_i).$$

Newton's third law states that the forces between bodies are equal in magnitude and opposite in direction, i.e. 'action = reaction'. This means that if there are no forces which act on the whole system (such as gravity), only interaction forces between particles, then momentum should always be conserved. Thus, conservation of momentum is generally a more useful test than energy conservation. However, as 'action = reaction' holds only for forces, we cannot invoke such a law for torques when the distance between the force point and the center of mass is different for the two bodies. The total angular momentum should be conserved

unconditionally; it is the sum of all orbital and intrinsic angular momenta,

$$\mathbf{L}^{\text{total}} = \mathbf{L}^{\text{total,o}} + \mathbf{L}^{\text{total,i}}.$$

For all particles the orbital angular momenta $\mathbf{L}_k^{\text{o}}$, cross products of the centers of mass $\mathbf{r}_k$ and the momenta $\mathbf{p}_k$, are summed:

$$\mathbf{L}^{\text{total,o}} = \sum_k \mathbf{L}_k^{\text{o}} = \sum_{k=1}^{n} \mathbf{r}_k \times \mathbf{p}_k = \sum_k m_k \mathbf{r}_k \times \mathbf{v}_k.$$

The intrinsic angular momenta ('spin') are the products of the tensors of the moments of inertia $\mathbf{J}_k$ and the angular velocities $\boldsymbol{\omega}_k$:

$$\mathbf{L}^{\text{total,i}} = \sum_k \mathbf{L}_k^{\text{i}} = \sum_k \mathbf{J}_k \boldsymbol{\omega}_k.$$

If angular momenta are to be compared for equality at times t_1 and t_2 (provided no external torques act on the system), the computation has to be done in the same coordinate system. Equality should also be taken in the numerical sense, i.e. with finite precision; see Chapter 2, particularly the section on relative and rounding errors. While many exercises in mechanics favor center-of-mass calculations for multi-body systems, in terms of numerical computation these offer no advantage. Newton's first law states that if the sum of forces acting on a body is zero, the body's velocity will be constant: either it stays at rest or it will move in a straight line with constant velocity. In simulations, initial conditions that should lead to zero velocity may actually experience motion due to small noise terms (which arise, for instance, from using a finite time-step or from oscillatory motion generated by insufficiently damped interaction forces). For example, a block on a slope which, according to analytical calculations, should keep its position may slide downhill in a simulation; see § 3.1.

1.8   Further reading

A readable introduction to mechanics is the book by R. D. Gregory [38]. Merkin et al. [33] give an overview of mathematical stability theory which is applicable to mechanical systems. A not-too-difficult introduction to Hamiltonian systems and geometrical integration is the book by Leimkuhler and Reich [15]. An extensive analysis of the frictional oscillator and further references can be found in the article by Kunze [23]. Solution methods for various nonlinear oscillations are discussed in Mickens's book [39]. For a recent monograph on Newton–Euler dynamics, see Ardema [40]. Resonance phenomena in nonlinear systems are treated by Manevich and Manevich [41].

Exercises

1.1 Rotations and complex numbers
    a) Compute the eigenvalues of the two-dimensional rotation matrix A^ϕ of Equation (1.9).

b) Show that multiplication of the two-dimensional vector $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ by the matrix A^ϕ is equivalent to rotation of $\mathbf{v}$ by angle ϕ.

1.2 Quaternions

a) *Quaternion product.* Derive the rule for quaternion products, Equation (1.92), from the definitions in (1.44)–(1.47).

b) *Real representation of quaternion basis elements.* Show that the real matrices

$$
B_{\mathbb{1}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad
B_{\mathsf{I}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},
$$

$$
B_{\mathsf{J}} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \qquad
B_{\mathsf{K}} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}
$$

(1.141)

satisfy the same commutativity relations as $\mathbb{1}$, I, J and K in Equations (1.44) and (1.45).

c) *Complex representation for element quaternions.* Show that the complex matrices

$$
B_{\mathbb{1}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad
B_{\mathsf{I}} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},
$$

$$
B_{\mathsf{J}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad
B_{\mathsf{K}} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}
$$

(1.142)

satisfy the same commutativity relations as $\mathbb{1}$, I, J and K in Equations (1.44) and (1.45).

d) In quantum physics, to describe objects with multiples of spin $\frac{1}{2}$, the Pauli matrices

$$
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad
\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad
\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
$$

(1.143)

are used. Compute the eigenvalues of the matrices (1.143) and the 2×2 matrices in (1.142). What is different?

e) Program the elementary operations for quaternions (multiplication, conjugation, inversion) as MATLAB® functions.

1.3 For the undamped case of the driven linear oscillator (1.101), i.e.

$$
\ddot{x} + \omega_0^2 x = f_0/m \exp(i\omega t),
$$

(1.144)

check for yourself that at resonance $\omega_0 = \omega$, not only $x(t) \propto \sin(i\omega t)$ is a solution but also $x(t) \propto t \sin(i\omega t)$, which means that although mathematically the resonance amplitude can become infinite according to Equation (1.105), its growth is only linear in time, so that the time needed to reach the infinite amplitude is also infinite.

References

[1] M. Nagasawa, *Schrödinger Equations and Diffusion Theory*. Monographs in Mathematics, Birkhäuser, 1993.

[2] G. Emch and C. Liu, *The Logic of Thermo-Statistical Physics*. Physics and Astronomy Online Library, Springer, 2002.

[3] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3rd ed. Pearson, 2001.

[4] M. P. Allen and D. Tildesley, *Computer Simulation of Liquids*. Oxford University Press, 1987.

[5] J. Myers, *Handbook of Equations for Mass and Area Properties of Various Geometrical Shapes*. NAVWEPS Report 7827, U.S. Naval Ordnance Test Station, 1962.

[6] J. Wittenburg, *Dynamics of Systems of Rigid Bodies*. Teubner, 1977.

[7] D. Greenwood, *Advanced Dynamics*. Cambridge University Press, 2006.

[8] H. Corben and P. Stehle, *Classical Mechanics*. Wiley, 1960.

[9] W. Benenson, J. Harris, H. Stöcker, and H. Lutz (eds.), *Handbook of Physics*. Springer, 2002.

[10] J. Hartog, "LXXIII. Forced vibrations with combined viscous and coulomb damping", *Philosophical Magazine Series 7*, vol. 9, no. 59, pp. 801–817, 1930.

[11] R. Reissig, "Erzwungene Schwingungen mit zäher und trockener Reibung", *Mathematische Nachrichten*, vol. 11, no. 6, pp. 345–384, 1954.

[12] J. Knudsen and P. Hjorth, *Elements of Newtonian Mechanics: Including Nonlinear Dynamics*, 3rd ed. Advanced Texts in Physics, Springer, 2000.

[13] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed. Graduate Texts in Mathematics, Springer, 1989.

[14] S. Zdravkovska and P. Duren, *Golden Years of Moscow Mathematics*. History of Mathematics, American Mathematical Society, 2007.

[15] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2004.

[16] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed. Springer, 2006.

[17] S.-H. Chong, M. Otsuki, and H. Hayakawa, "Generalized Green-Kubo relation and integral fluctuation theorem for driven dissipative systems without microscopic time reversibility", *Physical Review E*, vol. 81, 041130, 2010.

[18] A. Filippov, *Differential Equations with Discontinuous Righthand Sides*. Mathematics and its Applications, Springer, 1988.

[19] K. Johnson, *Contact Mechanics*. Cambridge University Press, 1987.

[20] W. C. Cheng, J. Chen, and H.-G. Matuttis, "Granular acoustics of polyhedral particles", *AIP Conference Proceedings*, vol. 1542, pp. 567 –570, 2013.

[21] S. A. E. Shourbagy, S. Okeda, and H. G. Matuttis, "Acoustic of sound propagation in granular materials in one, two, and three dimensions", *Journal of the Physical Society of Japan*, vol. 77, no. 3, article 034606, 2008.

[22] N. Fletcher and T. Rossing, *The Physics of Musical Instruments*. Springer, 1998.

[23] M. Kunze, "Rigorous methods and numerical results for dry friction problems", in *Applied Nonlinear Dynamics and Chaos of Mechanical Systems with Discontinuities* (M. Wiercigroch and B. De Kraker, eds.), World Scientific Series on Nonlinear Science Series A, pp. 207–235, World Scientific, 2000.

[24] D. Knuth, *The Art of Computer Programming, Volumes 1–4A*. Addison-Wesley, 2011.

[25] Y. Tokura and H. Maekawa, "Direct numerical simulation of impinging shock wave/transitional boundary layer interaction with separation flow", *Journal of Fluid Science and Technology*, vol. 6, no. 5, pp. 765–779, 2011.

[26] L. Bunimovich, "On the ergodic properties of nowhere dispersing billiards", *Communications in Mathematical Physics*, vol. 65, pp. 295–312, 1979.

[27] S. T. Maddison, J. R. Murray, and J. J. Monaghan, "SPH simulations of accretion disks and narrow rings", *Publications of the Astronomical Society of Australia*, vol. 13, pp. 66–70, 1996.

[28] R. Speith and W. Kley, "Stability of the viscously spreading ring", *Astronomy and Astrophysics*, vol. 399, no. 2, pp. 395–407, 2003.

[29] T. Shinbrot, C. Grebogi, J. Widsom, and J. A. Yorke, "Chaos in a double pendulum", *American Journal of Physics*, vol. 60, pp. 491–499, 1992.

[30] E. N. Lorenz, "Deterministic nonperiodic flow", *Journal of the Atmospheric Sciences*, vol. 20, pp. 130–141, 1963.

[31] I. N. Bronshtein, K. A. Semendyayev, G. Musiol, and H. Muehlig, *Handbook of Mathematics*, 5th ed. Springer, 2007.

[32] F. R. Gantmaher, *Lectures in Analytical Mechanics*. MIR Publishers, 1970.

[33] D. Merkin, F. Afagh, and A. Smirnov, *Introduction to the Theory of Stability*. Texts in Applied Mathematics, Springer, 1997.

[34] W. Greiner, *Classical Mechanics: Systems of Particles and Hamiltonian Dynamics*. Classical Theoretical Physics, Springer, 2010.

[35] L. Poinsot, *Théorie nouvelle de la rotation des corps*. Bachelier, 1851.

[36] J. P. Vinti, "Conservation laws and Liapounov stability of the free rotation of a rigid body", *Celestial Mechanics*, vol. 1, pp. 59–71, 1969.

[37] M. H. Kaplan, *Modern Spacecraft Dynamics & Control*. Wiley, 1976.

[38] R. Gregory, *Classical Mechanics*. Cambridge University Press, 2006.

[39] R. Mickens, *Truly Nonlinear Oscillations: Harmonic Balance, Parameter Expansions, Iteration, and Averaging Methods*. World Scientific, 2010.

[40] M. Ardema, *Newton-Euler Dynamics*. Springer, 2006.

[41] A. Manevich and L. Manevich, *The Mechanics of Nonlinear Systems with Internal Resonances*. Imperial College Press, 2005.