7 The nonlinearity of natural processes: the case of the pendulum

The nonlinear pendulum (Model 3 without friction, and Model 3′ with friction)

A closer analysis of the dynamics of the pendulum, still starting from Newton's law, leads to a different formulation of the equation of motion, less rough in its approximation and more articulated in its details. Ignoring the approximation introduced in Chapter 5 regarding the consideration of small oscillations only, we have, depending on whether we do or do not take friction into account, equation (7.1) and (7.2) respectively:

$$ml\ddot{\vartheta}(t) + mg \sin \vartheta(t) = 0 \tag{7.1}$$

$$ml\ddot{\vartheta}(t) + kl\dot{\vartheta}(t) + mg \sin \vartheta(t) = 0 \tag{7.2}$$

The similarity between equations (7.1) and (7.2) and between equations (7.2) and (5.4) is immediately apparent. The only difference, in both cases, is that the component of the force of gravity that recalls the pendulum towards the position of equilibrium, when the pendulum is distanced from it, represented by the last terms in the sums in the first members, is no longer proportional to the angle that the pendulum forms with the vertical, but to its sine. It is actually the presence of the term $\sin \vartheta(t)$, in this case, that makes the equation nonlinear with respect to the unknown function $\vartheta(t)$.

In the case that we are discussing, Figure 5.1 that describes the geometry of the system is transformed into Figure 7.1.

If we replace the function $\sin \vartheta$, that appears in (7.1) and (7.2) with its development in a Taylor series around the value $\vartheta = 0$, i.e.

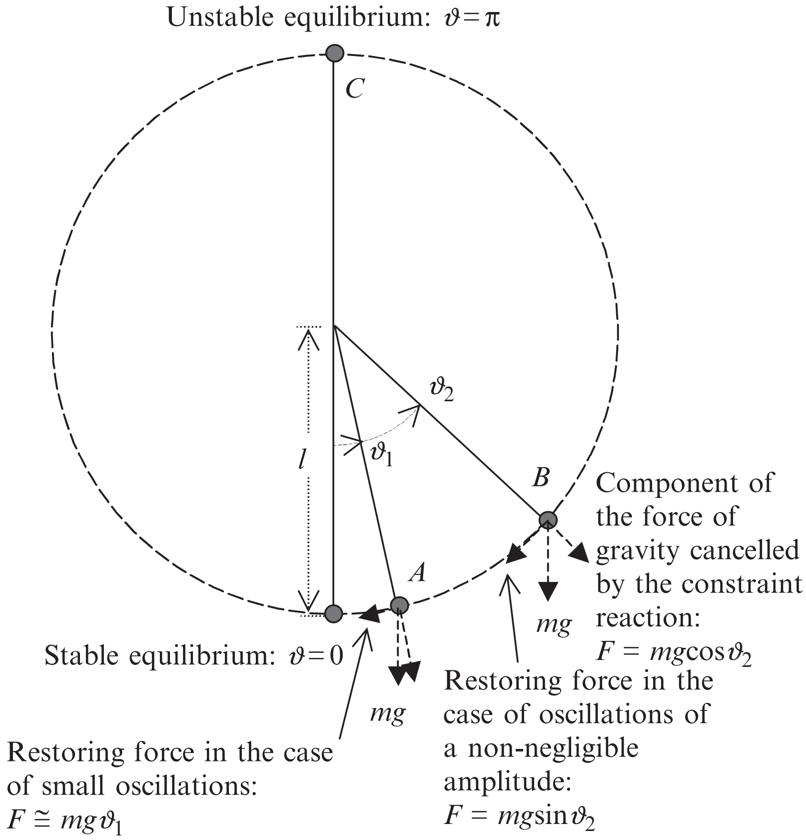

Figure 7.1 Diagram of a pendulum that makes oscillations of a non-negligible amplitude.

$$\sin(\vartheta) = \vartheta - \frac{\vartheta^3}{3!} + \frac{\vartheta^5}{5!} - \frac{\vartheta^7}{7!} + \ldots$$

we see that for small values of ϑ, i.e. for oscillations of an amplitude close to 0, we can ignore the terms of an order higher than first, which become increasingly smaller as the exponents that appear there increase. Equations (7.1) and (7.2) can thus be traced back to (5.2) and (5.4) respectively, i.e. they can be linearized. In this way, for small angles, the linear case appears to be a limit case of the general case.

Non-integrability, in general, of nonlinear equations

As we have said, while it is always possible to solve a linear differential equation, a nonlinear differential equation can only be solved in exact terms, i.e. without using approximations, in a limited number of cases only: the list of functions in which the solutions are usually expressed (polynomials, exponentials, etc.) is too limited to be able to adapt itself to the huge variety of nonlinear differential equations that we could be faced with in practice. Often, the attempt to write general solutions leads us to use integrations by series or integral functions, but methods of this nature are difficult to treat, are not always effective and can rarely be read intuitively. We need to focus, therefore, on other methods to qualitatively describe a phenomenon, after the differential equation has been formulated, methods that do not require the calculation of the general solution of the equation itself or methods that give an approximated solution (Jordan and Smith, 1977; Bender and Orszag, 1978). Each individual case should be studied in its own right; there are no general methods that can be applied to all the situations that we meet in models.

Given the historical importance of the problem of oscillations, and given the relative simplicity of the model that we are using as an example, we will see, for this case only, how it is possible to get around the difficulty of the lack of explicitly exact solutions[25] and make, in any event, several observations on the dynamics of the system.

We would like to immediately clarify that, even though we do not obtain exact solutions, it is still always possible to integrate (7.1) and (7.2) numerically, obtaining approximated solutions with arbitrary precision, at least in principle, for any set of initial conditions. What we are interested in is not so much giving life to a model, whose evolution is followed step by step (which we will do, however, for the logistic map in chapters 21 and 22), but rather in being able to formulate some general predictions on the future state of the system, at least where possible, even without a formula that provides the exact evolutionary trajectory of the system in question.

We will return to this point shortly. Before embarking on a discussion of oscillating systems, however, in chapter 8 that follows, we would like to make a digression of a more technical nature, where we will better clarify the two concepts that we have already used on bases that are little more than intuitive: the concepts of dynamical system and of phase space.

[25] Technically, the solution of (7.1) and (7.2) is ascribable to the calculation of so-called elliptic integrals, a class of integrals that is frequently applied in mathematical physics, that cannot be expressed by means of algebraic, logarithmic or circular functions, but that can only be calculated approximately, using an integration by series.

8 Dynamical systems and the phase space

What we mean by dynamical system

Let us reconsider a point that was introduced in chapter 2, but specifying the terms better: we want to define what is meant by dynamical system. When we describe the evolution of a system we use several state variables, i.e. a set of magnitudes, chosen appropriately, functions of time, the values of which define everything we know about the system *completely* and *unambiguously*; let us indicate them, just to be clear, with $x_i(t)$. Thus the evolution of a system is given by the evolution of the set of n state variables, each of which evolves according to a specific deterministic law. In this sense, we can speak of a dynamical system, identifying a system of this nature with the system of equations that defines the evolution of all of the state variables.

A dynamical system, therefore, is defined by a set of state variables and by a system of differential equations, of the following type:

$$\frac{d}{dt}x_i = F_i(x_1, \ldots, x_n, t) \quad i = 1, \ldots, n \tag{8.1}$$

In this case the dynamical system in continuous time is called a flow. Alternatively it can be defined by a set of state variables and by a system of finite difference equations, which make the system evolve at discrete time intervals:

$$x_i(k+1) = f_i(x_1(k), \ldots, x_n(k)) \quad i = 1, \ldots, n \tag{8.1$'$}$$

In this second case, the value of the variable x_i at instant $k + 1$ depends on that of all of the n variables at instant k; a dynamical system defined thus is called a map (occasionally also cascade) (Ansov and Arnold, 1988).

A mathematical model is usually written in the form of a dynamical system, i.e. in the form of (8.1) or (8.1′). In addition to these, there are also other forms in which dynamical systems and models can be formulated: for example, cellular automata, which are essentially dynamical systems in which, similar to finite difference equations, where time is considered discrete, the set of values that the state variables can assume is also assumed to be discrete (Wolfram, 1994), or in the form of partial derivative equations or integral equations.

Speaking in general terms, as we have already observed in Chapter 7, rather rarely can we integrate a dynamical system such as (8.1) analytically, obtaining a law that describes the evolution of the set of the state variables. The set of all of the possible dynamical systems (8.1) is in fact much more vast than the set of elementary functions such as sine, cosine, exponential, etc., that are used, in various ways, to integrate equations and to provide solutions expressed by means of a single formula (this is called at times 'in a closed form') (Casti, 2000).

We almost always need to use numerical integrations that transform the continuous form (8.1) into the discrete form (8.1′), or we need to use mathematical tools that enable us to study the qualitative character of the dynamical system's evolution (8.1). This has given rise to a sector of modern mathematics expressly dedicated to the study of dynamical systems and to the development of specific techniques for this purpose.

The phase space

The phase space (also called the space of states) is one of the most important tools used to represent the evolution of dynamical systems and to provide the main qualitative characteristics of such, in particular where the complete integration of a system is impossible. It is an abstract space, made up of the set of all of the possible values of the n state variables $x_i(t)$ that describe the system: for two variables the space is reduced to a plane with a pair of coordinate axes, one for $x_1(t)$, the other for $x_2(t)$; for three variables we would have a three-dimensional space and a set of three Cartesian axes, and so on for higher dimensions. In the case of the oscillation of a pendulum, for example, two state variables are needed to define the state and the evolution of the system: the angle $\vartheta(t)$ and the angular speed $\dot{\vartheta}(t)$.

Everything we know about the state of a dynamical system is represented by the values of the n coordinates in the phase space: knowing the values of the state variables is like knowing the values (dependent on time) of the coordinates obtained 'projecting' an object that we call a dynamical system on each of the n axes. The evolution of the system is given by the evolution of these variables in their entirety, in the sense that the law of dynamics (the law of motion) that defines the evolution of the system also defines the evolution over time of each of the n state variables. All that we need to know are the coordinates of the starting point, i.e. the initial values of the state variables. At times, however, it is more useful and easier to do the contrary, i.e. describe the system by obtaining its law of evolution from the law of evolution of each of the state variables considered individually. We should also add that a dynamical system is deterministic if the law of dynamics is such as to generate a single state consequent to a given state; on the other hand, it is stochastic, or random, if, consequent to a given state, there are more possible states from among which the dynamical system, in some way, can choose according to a probability distribution.

The phase space is a useful tool to represent the evolution of a dynamical system. Let us clarify this point with some basic examples. Let us consider a very simple case of dynamics, the simplest of all: a material point in uniform rectilinear motion in a direction indicated by a sole spatial coordinate x (Figure 8.1). In this case, (8.1) becomes simply: $\dot{x} = \text{constant}$.

The state of this system is characterized by a single spatial magnitude: $x(t)$, the coordinate of the point as a function of time (hereinafter, for the sake of simplicity, we will not write the explicit dependence of x on time). Uniform rectilinear motion means that the increase of x starting from an initial value x_0, i.e. the distance travelled, is directly proportional to the time t elapsed from the moment in which the mobile point was in x_0. We represent the dynamics of the point on a set of three Cartesian axes, in which the magnitudes x (the position), t (time) and $\dot{x}$ (speed) are shown on

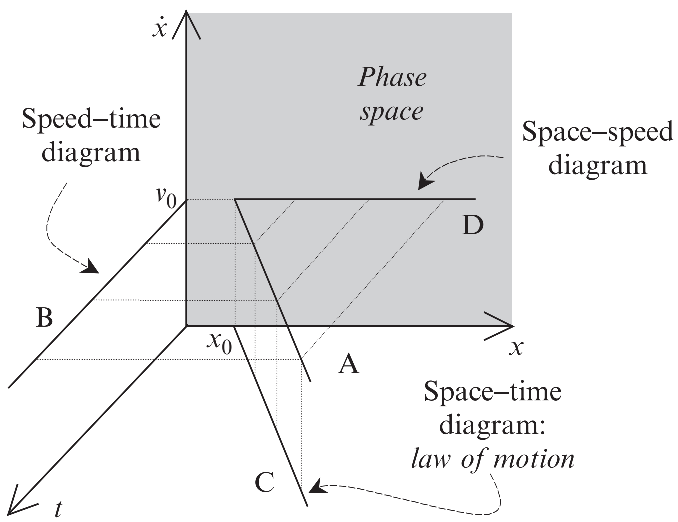

Figure 8.1 Dynamics of uniform rectilinear motion represented in a three-dimensional space x, $\dot{x}$, t and its projections on three planes.

the coordinate axes, as represented in Figure 8.1. In a case such as this, in which the motion is described by a sole coordinate, the plane $x\dot{x}$ is, precisely, the phase space (with just two dimensions: the phase plane).

It is then possible to consider the three diagrams that describe the uniform rectilinear motion: the space–time diagram (curve C in Figure 8.1), the speed–time diagram (curve B) and the space–speed diagram (curve D), as if they originated projecting a curve (curve A) on three planes xt, $\dot{x}t$ and $x\dot{x}$ in the three-dimensional space $x\dot{x}t$. It is immediately apparent that, in this case, this curve can only be a straight line, straight line A; a straight line, in actual fact, is the only curve that, projected on any plane, always and invariably gives a straight line.[26]

The same way of representing dynamics, applied to a case of motion that accelerates steadily with respect to a single variable x, gives Figure 8.2. Again in this case, in which (8.1) becomes $\dot{x} = kt$, the trend of the variables described by some of the known formulae of kinematics is immediately apparent.[27]

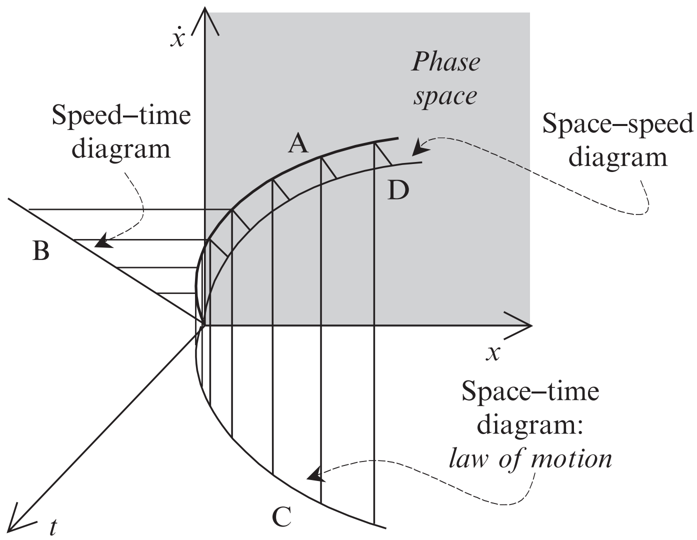

Figure 8.2 Dynamics of uniform accelerated motion represented in a three-dimensional space x, $\dot{x}$, t and its projections on three planes.

[26] More strictly speaking, we would like to remind readers that the equation that defines uniform rectilinear motion is $\ddot{x} = 0$; its integration is immediate and provides some of the laws of kinematics. In fact, a first integration with respect to time gives $\dot{x} = v_0$, the straight line B in Figure 8.1. A further integration gives $x = v_0 t + x_0$, curve C (law of motion). The speed, constant over time, is also constant in space, from which we obtain the horizontal straight line D in Figure 8.1 (space–speed diagram in the plane of phases). v_0 and x_0, respectively initial speed and initial coordinate, appear as constants of integration.

[27] The law of motion, in this case, is $\ddot{x} = a$, with a constant. By integrating once with respect to time and, for the sake of simplicity, assuming that the integration constants, i.e. the initial coordinate and the initial speed are nil, we have $\dot{x} = at$, curve B of Figure 8.2 (speed–

The usefulness of the phase space with respect to a space–time (the law of motion) or speed–time diagram lies in the fact that in the phase space the variables that appear are not represented as a function of time; time is only indirectly present, as the independent variable with respect to which space is derived.

This enables us to obtain a representation of the system's states in what we could call an atemporal way, that is to say in a perspective in which time does not play the primary role of the system's 'motor element', but rather the secondary role as a reference for the speed at which a spatial magnitude varies. In other words, the system is represented only with respect to the state variables (a bit like saying 'with respect to itself'), without the introduction of variables, such as time, that do not characterize the system itself, and in this sense, are not strictly necessary.

Therefore, the information shown in a phase space diagram is better, because it contains as much information as that provided by two space–time and speed–time diagrams, but it is expressed more succinctly in a single diagram. Obviously, this only applies to autonomous differential equations, in which time is not expressly present; in the case of non-autonomous systems, on the contrary, time is one of the variables which define the phase space.

The trajectories traced by a system in the phase space are commonly called *orbits*.

To conclude this digression into the phase space, we can add an initial definition of the concept of the stability of a system, a concept that we already briefly mentioned in our discussion of the pendulum, using the phase space as a tool to represent the dynamics; we will be discussing the concept of stability in further detail in Chapters 17–19. We will be looking at the concept of a system's stability again in Chapter 12, when we discuss the evolution of a system of two interacting populations.

If a system is in some way distanced from equilibrium, as a result of said movement, it may demonstrate different dynamics, depending on the type of law that governs it: (1) the system may evolve, remaining close to the abandoned equilibrium configuration, which, in this case is a stable equilibrium; (2) the system may tend to return to the abandoned equilibrium, which, in this case is a asymptotically stable equilibrium; (3) the system may distance itself further still from the equilibrium, which in this case is known as unstable equilibrium.

time diagram); by integrating a second time, we have $x = at^2/2$, curve C (law of motion). By eliminating time from the last two equations, we obtain the expression of speed with respect to space, i.e. the trajectory in the phase space: $\dot{x} = \sqrt{2ax}$, i.e. curve D (space–speed diagram).

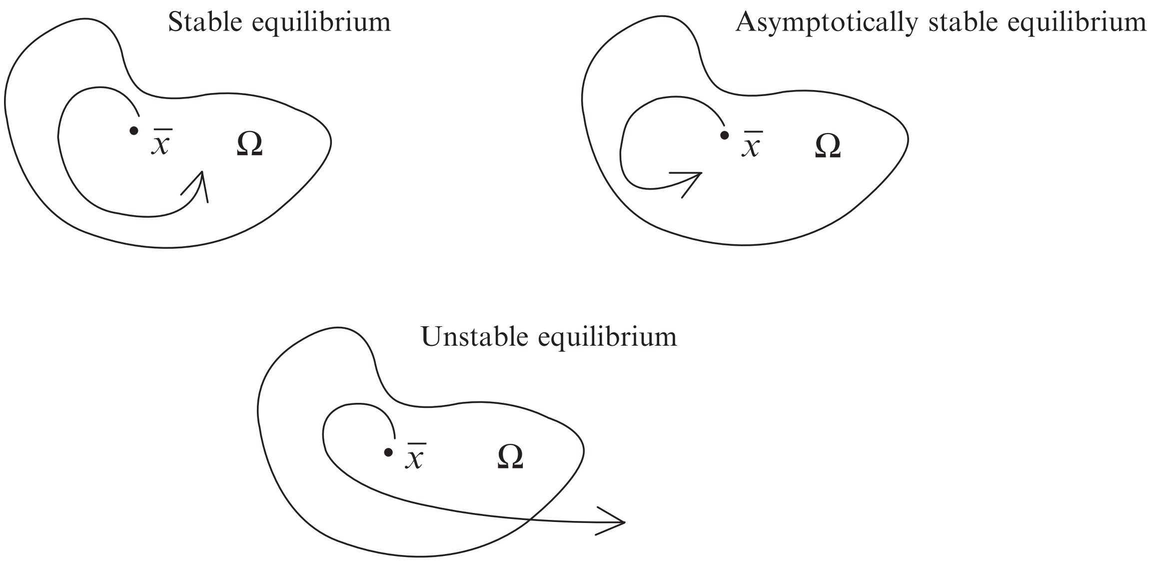

Figure 8.3 Depiction of stable, asymptotically stable and unstable equilibrium.

More precisely, we can indicate the set of the values of the coordinates of a point of equilibrium as $\bar{x}$: if the system's orbit remains indefinitely confined to a neighbourhood Ω of $\bar{x}$ (Figure 8.3), then $\bar{x}$ is a point of stable equilibrium (this is the case, for example, of the *non-damped* oscillations of a pendulum around the rest position). If on the other hand, as time tends towards infinity, the successive states of the system always tend increasingly towards $\bar{x}$, then $\bar{x}$ is an asymptotically stable point of equilibrium (this is the case, for example, of the dampened oscillations of a pendulum around a position of equilibrium). Lastly, if no neighbourhood of $\bar{x}$ can be found that satisfies what has been defined, then $\bar{x}$ is an unstable point of equilibrium of the system (this is the case, for example of a pendulum placed vertically, with the mass above the centre of rotation, as in case C in Figure 7.1, in which $\vartheta = \pi$).

Oscillatory dynamics represented in the phase space

Let us now return to the dynamics of the pendulum according to the linear model (Model 1: small oscillations without friction) that we described in Chapter 4 (hereinafter we will refer to the linear model by its proper name in mathematical physics: 'harmonic oscillator'). The considerations that we will make are quite technical: on one hand they can be considered as an example of the use of the concepts that we have introduced so far (the phase space, stable/unstable equilibrium, etc.); on the other, they actually enable us to introduce other concepts (attractor, unstable trajectories, etc.) that we will also find in cases of less basic dynamics.

As we know, the solution of the equation of motion (5.2) is the law of motion (5.3), in which $\omega^2 = g/l$. They are both shown below:

$$ml\ddot{\vartheta}(t) = -mg\vartheta(t)$$

$$\vartheta(t) = \vartheta_0 \cos \omega t$$

Equation (5.2) is a second order equation that can be either fully (i.e. twice) integrated with respect to time or just once. By integrating once only, with some intermediate steps, we thus obtain a first order differential equation, a so-called first integral of motion:[28]

$$\frac{1}{2}ml^2\dot{\vartheta}^2 + \frac{1}{2}mgl\vartheta^2 = \text{constant} \tag{8.2}$$

The first integral (8.2) expresses the well-known fact that during the oscillation of a system, there is one magnitude that remains constant over time, and is therefore called a 'constant of motion': we will indicate it with E, and we will call it the system's total energy. In the case in which the harmonic oscillator is in actual fact the linear pendulum we discussed earlier, we can demonstrate that the two terms that are summed to the first member of (8.2) are, respectively, the kinetic energy and the potential energy of the pendulum, and the total energy is simply the mechanical energy. A system in which mechanical energy is a constant of motion is called *conservative*.[29]

[28] More precisely, given a differential equation of order $n \geq 2$, all relationships between the independent variable, the unknown function and its first $n-1$ derivatives are called first integrals. The study of the first integrals of the equations of motion enables us to identify constants of motion that can be as fundamental to the description of the phenomena, as the complete solution of equations of motion. It is precisely from the examination of first integrals that in the eighteenth century, the formulation of important general principles (or rather theorems) for the conservation of mechanical systems was achieved, such as the theorem of the conservation of mechanical energy, of linear momentum and of angular momentum, a theorem that played a fundamental role in the birth and the development of deterministic concepts in the science of the era (see Chapter 3).

[29] The fact that a constant of motion, mechanical energy, *exists* is the well-known theorem of the conservation of mechanical energy of a material point, one of the greatest conquests of mathematical physics and, in general, of scientific thought of the eighteenth century. This *theorem* represents a specific case of the general *principle* of energy conservation (if non-conservative forces are acting), fundamental to all branches of physics. We observe that we have discovered that in oscillations, a constant of motion *exists*, without fully integrating the equation of motion (5.2), therefore without calculating the real law of motion (i.e. the law that links space to time).

In the phase space defined by the variables ϑ and $\dot{\vartheta}$ (in this case, therefore, it is a phase plane), the curves represented by (8.2), as E varies, constitute a set of ellipses, the centre of all of which is at the origin of the coordinate axes and the lengths of whose semiaxes are, respectively, $\sqrt{2E}$ and $\sqrt{2E}/\omega$. Therefore, the dimensions of the orbit in the phase space depend on energy E. In Figure 8.4 we show several different values of E.

Each of the ellipses traced is the set of the points in the phase space that define states of equal total energy. The arrows indicate the direction in which the orbits are travelled with the passing of time.

As the system travels along one of the indicated orbits, the energy of the pendulum is transformed from kinetic to potential, then from potential to kinetic, then from kinetic to potential again and so on; the two forms of energy alternate periodically, maintaining their sum constant. As in the approximation considered, in which friction is not introduced, there is no dispersion of mechanical energy, because the latter is a constant of motion, the possible orbits are all isoenergetic and correspond to different quantities of total energy. Different orbits do not intersect each other, because a point of intersection, belonging to two different orbits, would result in two different values of the total energy (which is constant) for the same system.

Let us leave the harmonic oscillator and return to the linear pendulum subject to friction described in equation (5.4) (Model 2). With considerations similar to those described above, we discover, in this case, that the total mechanical energy (only the mechanical energy!) E is *not* maintained, but dissipates over time.[30] In this case the system is called dissipative; the

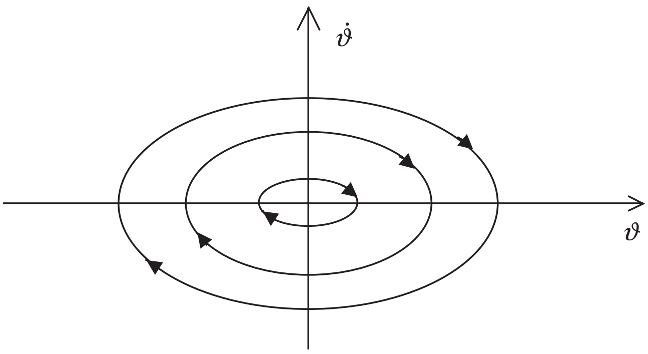

Figure 8.4 Representation of the motion of a harmonic oscillator in the phase space for different values of E.

[30] It must be absolutely clear to the reader that, when we speak of dissipated energy, we are only referring to mechanical energy (potential energy + kinetic energy), *not* to total energy (mechanical energy + thermal energy + energy of the electromagnetic field + mass energy + ...). According to current scientific vision and knowledge, the total energy of a closed system is *always* conserved in *any* process.

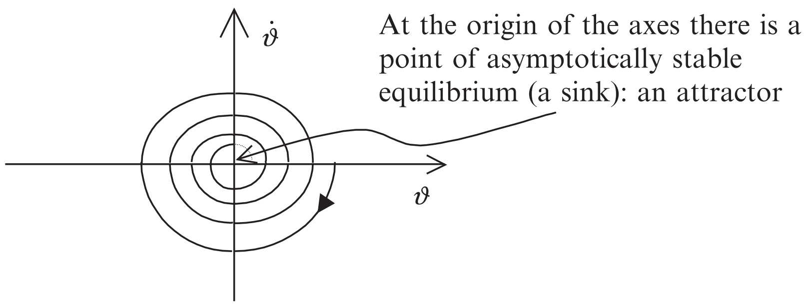

Figure 8.5 Representation of the motion of an oscillating system in the phase space that undergoes a damped oscillation in the presence of friction.

orbit in the phase space $\vartheta\dot{\vartheta}$ is now made up of a spiral with an asymptotically stable point of equilibrium at the origin of the axes (sometimes this point is called a 'sink'), i.e. at the angle $\vartheta = 0$ and with the pendulum still: $\dot{\vartheta} = 0$ (Figure 8.5).

We have therefore met a first, very basic example of a dissipative system, whose orbit in the phase space is an open curve. We will repeatedly come across dissipative systems characterized by non-periodic orbits in the second part.

Let us now consider the motion of the pendulum that makes oscillations of large amplitude in the simpler form described by (7.1), i.e. ignoring friction (undamped oscillations: Model 3). Equation (7.1) can be rewritten as follows:

$$m\ddot{\vartheta} + m\omega^2 \sin\vartheta = 0 \tag{8.3}$$

By integrating (8.3) once with respect to time, as we did with (5.2), we discover the existence of a magnitude that is maintained constant over time. Here again, as for (8.2), we are dealing with total mechanical energy E:

$$\frac{1}{2}ml^2\dot{\vartheta}^2 + mgl(1 - \cos\vartheta) = E \tag{8.4}$$

Just as in (8.2), (8.4) which represents the general case for any angles ϑ, the first addend to the first member is the pendulum's kinetic energy, while the second addend is the potential energy.

Oscillations of small amplitudes mean small energies and, as we have said, in this case (8.2) gives elliptic orbits. For oscillations of amplitudes that cannot be considered 'small', i.e. if linear approximation is not acceptable, the dynamics of the pendulum is no longer that provided by the model of the harmonic oscillator (5.2) and the orbits in the phase space are curves other than ellipses (8.2). We develop the term $\cos\vartheta$ in Taylor series around $\vartheta = 0$:

$$\cos\vartheta = 1 - \frac{\vartheta^2}{2!} + \frac{\vartheta^4}{4!} - \frac{\vartheta^6}{6!} + \ldots$$

If we ignore, for small ϑ, the terms above the second order, the equation (8.4) leads back to equation (8.2) and therefore we return to linear approximation. If, on the other hand, as the amplitude of oscillation ϑ increases, we progressively consider the subsequent terms of the series development of $\cos\vartheta$, the curves that (8.4) gives in the phase space defined by variables ϑ and $\dot{\vartheta}$ remain closed, but we get increasingly further away from the ellipse shape (8.2), with an increasingly marked difference as energy E grows. We can demonstrate that, in correspondence to a critical value of E, the orbits become closed curves formed by sinusoidal arcs in the upper half-plane and by sinusoidal arcs in the lower half-plane. For higher energy values, the orbits are not even closed curves: the pendulum no longer oscillates around a point of stable equilibrium, but rotates periodically.[31]

[31] Let us add some technical considerations that are not indispensable to the discussion presented in the text. Remembering that $\omega^2 = g/l$, we have from (8.4):

$$\dot{\vartheta}^2 = \frac{2E}{ml^2} - 2\omega^2 + 2\omega^2\cos\vartheta \tag{8.5}$$

which, for small values of ϑ, is similar to the form of the harmonic oscillator:

$$\dot{\vartheta}^2 \cong \frac{2E}{ml^2} - \omega^2\vartheta^2 \tag{8.5$'$}$$

which is simply (8.2) rewritten expressing $\dot{\vartheta}^2$.

As the mechanical energy E (positive) varies, (8.5) represents a set of orbits in the phase space defined by the variables ϑ and $\dot{\vartheta}$ (Figure 8.6):

1. For $0 < E < 2ml^2\omega^2$, we have the *closed orbits* cited in the text: (i) in the limit case of small values of the angle of oscillation ϑ (the case of the harmonic oscillator), they become the ellipses given by (8.5′); (ii) in the general case of large values of ϑ, they are the closed curves (deformed ellipses) given by (8.5).
2. For $E = 2ml^2\omega^2$, the critical value of the energy cited in the text above, we have the case of *closed orbits* composed of arcs belonging to the two cosine curves that intersect each other on the axis of the abscissas of Figure 8.6. This is the limit case that separates the previous case from the following case: it corresponds to the motion of a pendulum whose energy is ‘exactly’ equal to that which would allow it to reach the vertical position above the centre of oscillation (point C in Figure 1.3) with nil speed and in an infinite time (the total energy E is ‘exactly’ equal to the potential energy of the pendulum in equilibrium above the centre of rotation: $U(\vartheta) = 2mgl$).
3. For $E > 2ml^2\omega^2$, (8.5) represents a set of cosine curves, symmetrical in pairs with respect to the axis of the abscissas, shifted higher (or lower) with respect to those of the previous case, the higher the value of E, and with amplitudes that decrease with the entity of the

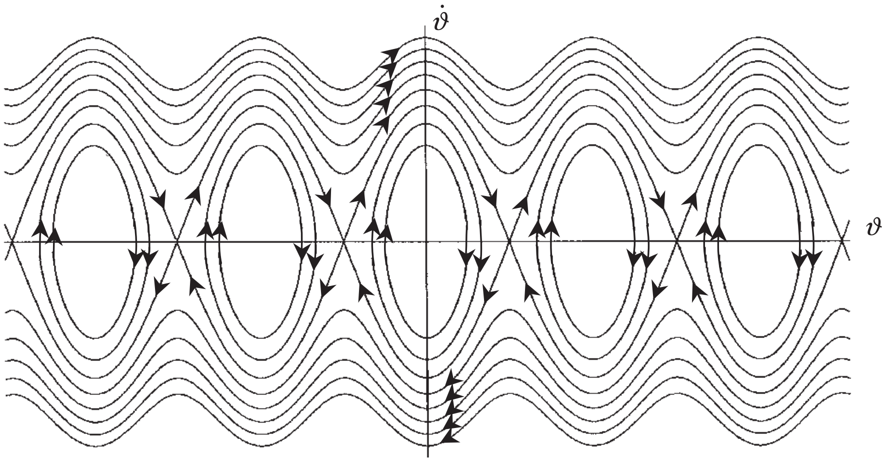

Figure 8.6 Orbits in the phase space of an undamped pendulum.

shifting. These are *open orbits* shown at the top (and bottom) of Figure 8.6: the pendulum has an energy E that is so great that it does not oscillate, but rotates periodically at a speed $\dot{\vartheta}$ that is not constant (the orbits are oscillating functions), but, on average, is greater the higher E is.