Chapter 2

Multiple variables

2.1 Two variables

2.1.1 First order coupled ODEs

While the position of a particle along a line can be represented by a single coordinate, its location on a two-dimensional plane requires two coordinates, say indicated by x_1 and x_2. In the absence of time-reversal symmetry, the generalization of Eq. (1.1.10) to two degrees of freedom is (setting $\mu = 1$ without loss of generality)

$$\dot{x}_1 = F_1(x_1, x_2), \qquad \text{and} \qquad \dot{x}_2 = F_2(x_1, x_2)\,. \tag{2.1.1}$$

Let us assume that $x_1 = x_2 = 0$ is a point of equilibrium (at which $F_1 = F_2 = 0$). Series expansions[1] of the force around this point then yield to the lowest order

$$F_1(x_1, x_2) = f_{11}x_1 + f_{12}x_2 + \cdots\,, \qquad \text{and} \qquad F_2(x_1, x_2) = f_{21}x_1 + f_{22}x_2 + \cdots\,. \tag{2.1.4}$$

Understanding the behavior of the system near $x_1 = x_2 = 0$ thus requires solving the pair of coupled first order ODEs

$$\begin{cases} \dot{x}_1 = f_{11}x_1 + f_{12}x_2 \\ \dot{x}_2 = f_{21}x_1 + f_{22}x_2 \end{cases}, \quad \Longrightarrow \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \Longrightarrow \quad \frac{d\vec{x}}{dt} = \mathbf{F} \cdot \vec{x}\,. \tag{2.1.5}$$

---

[1] The Taylor expansion of a function of two variables takes the form

$$\phi(x, y) = \phi_{00} + \phi_{10}x + \phi_{01}y + \frac{\phi_{20}}{2}x^2 + \phi_{11}xy + \frac{\phi_{02}}{2}y^2 + \cdots \equiv \sum_{m,n} \phi_{mn} \frac{x^m}{m!} \frac{y^n}{n!}\,, \tag{2.1.2}$$

with the coefficients obtained from mixed derivatives, as

$$\phi_{mn} = \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} \phi(x, y)\bigg|_{x=y=0}\,. \tag{2.1.3}$$

Note that the symbol $\partial/\partial x$ is used in place of d/dx, indicating partial derivatives of the function with respect to the variable x, when other variables of the function are held constant.

Note that the linear set of equations can be cast in the form of a 2×2 matrix acting on the column vector composed from x_1 and x_2.

To gain insight, let us first consider a particle sliding down a two-dimensional potential shaped like an ellipsoidal bowl. If we align the coordinates x_1 and x_2 to the axes of the ellipse, the expansion of the potential around its minimum at $x_1 = x_2 = 0$ reads

$$V(x_1, x_2) = k_1 \frac{x_1^2}{2} + k_2 \frac{x_2^2}{2} + \cdots \,, \tag{2.1.6}$$

where k_x and k_y are the inverse radii of curvature of the bowl. A particle that starts along one of the axes of ellipse, say from $(x_1 = x_1^0, x_2 = 0)$, will stay on the axis as the symmetry $x_2 \to -x_2$ does not select one direction over the other ($F_2 = -k_2 x_2 = 0$ for $x_2 = 0$), and proceeds according to

$$\dot{x}_1 = F_1 = -\frac{\partial V}{\partial x_1} = -k_1 x_1 \,, \quad \Longrightarrow \quad x_1(t) = x_1^0 e^{-k_1 t} \,. \tag{2.1.7}$$

A corresponding solution can be obtained for a particle that starts at $(x_1 = 0, x_2 = x_2^0)$ at $t = 0$. The general solution, starting from any initial point is the simple superposition of the two, decaying to zero as

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x_1^0 e^{-k_1 t} \\ x_2^0 e^{-k_2 t} \end{pmatrix} \,. \tag{2.1.8}$$

Let us suppose, however, that by some curious oversight we had chosen to align our coordinate system at 45° (or some other angle) with respect to the natural directions of the elliptical bowl. The new coordinates $\{x'_1, x'_2\}$, and the old ones are related by[2]

$$\begin{cases} x'_1 = \dfrac{x_1 + x_2}{\sqrt{2}} \\ x'_2 = \dfrac{-x_1 + x_2}{\sqrt{2}} \end{cases} , \qquad \text{and} \qquad \begin{cases} x_1 = \dfrac{x'_1 - x'_2}{\sqrt{2}} \\ x_2 = \dfrac{x'_1 + x'_2}{\sqrt{2}} \end{cases} . \tag{2.1.10}$$

In terms of the new coordinates, the potential energy is

$$V = \frac{k_1}{4} \left(x'_1 - x'_2\right)^2 + \frac{k_2}{4} \left(x'_1 + x'_2\right)^2 = \frac{k_1 + k_2}{4} \left(x'^2_1 + x'^2_2\right) - \left(\frac{k_1 - k_2}{2}\right) x'_1 x'_2 \,. \tag{2.1.11}$$

The corresponding equations of motion for gradient descent,

$$\begin{cases} \dot{x}'_1 = -\dfrac{\partial V}{\partial x'_1} = -\dfrac{k_1 + k_2}{2} x'_1 + \dfrac{k_1 - k_2}{2} x'_2 \\ \dot{x}'_2 = -\dfrac{\partial V}{\partial x'_2} = -\dfrac{k_1 + k_2}{2} x'_2 + \dfrac{k_1 - k_2}{2} x'_1 \end{cases} , \tag{2.1.12}$$

---

[2] For a rotation by an angle θ, we have

$$\begin{cases} x' = x \cos\theta + y \sin\theta \\ y' = -x \sin\theta + y \cos\theta \end{cases} , \qquad \text{and} \qquad \begin{cases} x = x' \cos\theta - y' \sin\theta \\ y = x' \sin\theta + y' \cos\theta \end{cases} . \tag{2.1.9}$$

are *coupled* to each other, and a may appear harder to solve. Naturally, from the original solution to the problem, it is easy to construct solutions to these equations as

$$\begin{cases} x_1'(t) = \dfrac{x_1(t)+x_2(t)}{\sqrt{2}} = \dfrac{1}{\sqrt{2}}\left[x_1^0 e^{-k_1 t} + x_2^0 e^{-k_2 t}\right] \\ x_2'(t) = \dfrac{x_2(t)-x_1(t)}{\sqrt{2}} = \dfrac{1}{\sqrt{2}}\left[x_2^0 e^{-k_2 t} - x_1^0 e^{-k_1 t}\right] \end{cases} . \tag{2.1.13}$$

However, the solutions in the case are not single exponentials, but superposition of two exponentials.

This example demonstrates that there could be a *'right way'* of looking at a system, and many possible *'wrong ways'* (rotated coordinates) of viewing it. The analysis and description of the system becomes much simpler if the right set of coordinates are used. Surprisingly, this is always possible for gradient descent in a linear system, and there is a way to find the right set of variables to describe the problem.

2.1.2 Gradient descent in two dimensions

The most general form of a quadratic potential in two dimensions, generalizing Eq. (2.1.6), is[3]

$$V(x_1, x_2) = V_0 + k_1\frac{x_1^2}{2} + k_2\frac{x_2^2}{2} + k_\times x_1 x_2 \,. \tag{2.1.14}$$

Gradient descent in such a potential leads to

$$\begin{cases} \dot{x}_1 = -\dfrac{\partial V}{\partial x_1} = -k_1 x_1 - k_\times x_2 \\ \dot{x}_2 = -\dfrac{\partial V}{\partial x_2} = -k_2 x_2 - k_\times x_1 \end{cases} , \quad \Longrightarrow \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = -\begin{pmatrix} k_1 & k_\times \\ k_\times & k_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} . \tag{2.1.15}$$

The analog of the direction in Eq. (2.1.7) along which the solution proceeds exponentially, is an eigenvector of the above matrix,

$$\mathbf{M} = -\begin{pmatrix} k_1 & k_\times \\ k_\times & k_2 \end{pmatrix} , \tag{2.1.16}$$

with the decay rate provided by the corresponding eigenvalue. In other words, we seek column vectors

$$\vec{e}_\pm \equiv \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad \text{such that} \quad \mathbf{M}\vec{e}_\pm = \lambda_\pm \vec{e}_\pm \,. \tag{2.1.17}$$

The indices ± are in anticipation of there being two directions and corresponding eigenvalues.

To obtain the eigenvalues, the equation is first rearranged as $(\mathbf{M} - \lambda\mathbf{1}) \cdot \vec{e} = \mathbf{0}$, where $\mathbf{1}$ is the unit matrix with ones along the diagonal and zeros elsewhere. For this homogenous

---

[3] Linear terms in x_1 and x_2 are absent, either because of an inversion symmetry $\vec{x} \to -\vec{x}$, or because we are interested in deviations from a stable equilibrium.

set of equations to have a non-zero answer, the *determinant* of the matrix of coefficients has to be zero, i.e.

$$\det(\mathbf{M} - \lambda \mathbf{1}) = 0.. \tag{2.1.18}$$

For our 2×2 matrix, this leads to a so-called *characteristic equation* that has the form

$$\lambda^2 - \mathrm{tr}\mathbf{M}\ \lambda + \det \mathbf{M} = 0\,, \quad \text{with} \quad \mathrm{tr}\mathbf{M} = k_1 + k_2 \quad \text{and} \quad \det \mathbf{M} = k_1 k_2 - k_\times^2\,. \tag{2.1.19}$$

It is good to recall that the sum of the two eigenvalues is equal to the trace of the matrix, while their product is the determinant of the matrix. Solving this quadratic equation gives

$$\lambda_\pm = -\frac{1}{2}\left[(k_1 + k_2) \pm \sqrt{(k_1 - k_2)^2 + 4k_\times^2}\right]\,. \tag{2.1.20}$$

Note that the quantity under the square root is strictly positive, indicating that both eigenvalues are real. For stable equilibrium, both eigenvalues should be negative, as positive eigenvalues will take the dynamics to infinity; this occurs for $k_\times^2 > k_1 k_2$ where $\det \mathbf{M} < 0$.

2.1.3 Beyond gradient descent

The reality of the eigenvalues in Eq. (2.1.20) is a consequence of the symmetry of the matrix, which is an inevitable consequence of gradient descent in a quadratic potential. However, even for more complicated potentials gradient descent (that $F_1 = \frac{\partial V}{\partial x_1}$ and $F_2 = \frac{\partial V}{\partial x_2}$) imposes the constraint

$$\frac{\partial F_1}{\partial x_2} = -\frac{\partial^2 V}{\partial x_2 \partial x_1} = -\frac{\partial^2 V}{\partial x_1 \partial x_2} = \frac{\partial F_2}{\partial x_1}\,, \tag{2.1.21}$$

since mixed partial derivatives can be taken in any order. If so, we may ask what is the outcome of the more general dynamics that does not satisfy the above constraint, e.g. for linearized equations such as in Eq. (2.1.5), where the matrix

$$\mathbf{F} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}\,, \tag{2.1.22}$$

is not symmetric, $f_{12} \neq f_{21}$?

As example, let us consider the following set of equations

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\gamma v - \omega_0^2 x \end{cases}, \quad \Longrightarrow \quad \begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix}\,. \tag{2.1.23}$$

Clearly this systems of two coupled first order equations is simply the damped harmonic oscillator of Eq. (1.4.8) in disguise. The eigenvalues of the asymmetric matrix are given by Eq. (1.4.10). Notably, for $\gamma < 2\omega_0$ the eigenvalues have an imaginary part indicating oscillatory behavior. In the limit $\gamma = 0$, the motion is undamped oscillation (time reversible) and conserves the energy function $E(x, v) = (v^2 + \omega_0^2 x^2)/2$.

2.1.4 Hamiltonian evolution

Conservation of energy is an important principle in physics, and it is useful to find a procedure to construct first order equations that conserve some function, say $H\left(x(t), p(t)\right)$. Setting $dH/dt = 0$, and using the chain rule, we need

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial p}\dot{p} = 0\,. \tag{2.1.24}$$

One way to ensure this condition is to set

$$\dot{x} = \frac{\partial H}{\partial p}\,, \qquad \text{and} \qquad \dot{p} = -\frac{\partial H}{\partial x}\,. \tag{2.1.25}$$

Clearly, Eqs. (2.1.23) for $\gamma = 0$ follow this structure with v playing the role of p. Indeed, the *Hamiltonian formulation* of classical equations of motion follow the structure of Eq. (2.1.24) and (2.1.25), with $H(x, p)$ as the total energy in terms of the coordinate x and its conjugate momentum p.

Indeed the most general pair of linear ODEs from Eq. (2.1.26) can be recast as

$$\dot{x}_1 = F_1(x_1, x_2) = -\frac{\partial V}{\partial x_1} + \frac{\partial H}{\partial x_2}, \qquad \text{and} \qquad \dot{x}_2 = F_2(x_1, x_2) = -\frac{\partial V}{\partial x_2} - \frac{\partial H}{\partial x_1}\,, \tag{2.1.26}$$

as superposition of gradient descent in the potential $V(x_1, x_2)$ with sliding along contours of constant $H(x_1, x_2)$.[4]

Recap

- A general pair of first order ODEs can be cast as gradient descent in a potential V and sliding along contours of constant H.
- The linearized equations can be cast as a 2×2 matrix, whose eigenvalues determine the exponential rates along the two eigendirections.
- Symmetric matrices, corresponding to gradient descent in a quadratic potential, have two real eigenvalues. The eigenvalues of an asymmetric matrix may or may not be complex, with complex eigenvalues indicative of oscillatory behavior.

---

[4] For future reference, note that given F_1 and F_2, the potentials V and H are solutions to

$$\begin{cases} \dfrac{\partial^2 V}{\partial {x_1}^2} + \dfrac{\partial^2 V}{\partial {x_2}^2} = \dfrac{\partial F_1}{\partial x_1} + \dfrac{\partial F_2}{\partial x_2} & \Rightarrow \quad \nabla^2 V = \nabla \cdot \vec{F} \\ \dfrac{\partial^2 H}{\partial {x_1}^2} + \dfrac{\partial^2 H}{\partial {x_2}^2} = \dfrac{\partial F_1}{\partial x_2} - \dfrac{\partial F_2}{\partial x_1} & \Rightarrow \quad \nabla^2 H = \nabla \times \vec{F} \end{cases}. \tag{2.1.27}$$