Encyclopedia of Mathematical Physics, pp. 465–482,
coordonnée par J.P. Françoise, G. Naber et T.S. Tsun, Elsevier (2006).

Multi-scale approaches

Annick LESNE

Laboratoire de Physique Théorique de la Matière Condensée, Case 121,
Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05, France
lesne@lptmc.jussieu.fr

1 Introduction: multiple-scale and multi-scale approaches

Multi-scale or more precisely *multiple-scale method* is a technique of perturbation theory based on the introduction of additional rescaled variables, say time variables, formally considered as independent variables and describing each a different time scale (for the sake of simplicity, I will mainly consider a dynamic framework and time scales; all can be transposed to spatial dependences and scales). It has been first developed to handle singular situations in which dynamic regimes of different characteristic scales coexist and intermingle in such a way that straightforward perturbation expansions are not uniformly convergent in time (hence of limited relevance and use) due to so-called *secular terms* growing unbounded with time; the freedom introduced together with the extra variables indeed allows to impose conditions preventing from these secular divergences and improving the convergence of the perturbation series. It yields a global perturbation solution describing *jointly* the behavior at small and large scales. This technique belongs to the far wider ranging class of *multi-scale approaches*; these can be divided in four main sub-classes:
1) *mean-field techniques* exploiting scale separation between fast and slow components of the dynamics. The influence of the slow variables onto the fast dynamics, if any, is treated in a decoupled way within a parametric approximation, allowing an adiabatic elimination of fast variables (§ 4);
2) *singular perturbations*, in which individual fast components ultimately give rise to slow trends and influence the large-scale features. Scale separation here breaks down at long times and multiple-scale method is then a method of choice (§ 2);
3) *matched expansions* when regimes of different scales succeed (boundary-layer singularity, § 5);
4) *renormalization techniques*, in systems exhibiting some kind of universality in the relations between their behaviors at different scales, e.g. scale invariance (§ 6).

I'll first present the principles of multiple-scale method, detail its technical implementation on simple abstract examples and cite some typical applications. Then I'll articulate this technique with more general multi-scale methods in a brief overview (§ 3). The range of multi-scale approaches and technical tools will be then illustrated and compared in the context of diffusion, Brownian motion and transport phenomena (§ 7).

2 Multiple-scale method: principles

2.1 Context: singular perturbations and secular divergences

Multiple-scale methods have been developed to handle situations in which the dynamics involves a small parameter ϵ (e.g. the ratio of the masses of different subsystems, the strength of an additional interaction, the amplitude of an applied field) directly controlling the separation between the different characteristic time-scales of the evolution and, specifically, such that *the behavior for*

$\epsilon = 0$ *is qualitatively different from the behavior for* ϵ *small* ($\epsilon \ll 1$ but finite); in other words, when a weak influence, of strength controlled by $\epsilon \ll 1$, does not only have weak consequences. Typically, this occurs when ϵ represents the strength of a weak coupling between otherwise independent sub-systems or when a vanishing value $\epsilon = 0$ changes a characteristic time, the sign of a friction coefficient, the order of the highest time derivative in case of ordinary differential equations (turning points), or the type of partial differential equations in case of spatially extended systems. Accordingly, a naive perturbative approach with respect to ϵ, i.e. an expansion taking as a basic approximation the behavior for $\epsilon = 0$, cannot bridge the qualitative gap with behaviors observed for $\epsilon > 0$. It thus fails to give a full account of the system evolution at all times: one speaks of *singular perturbation.*

An historical example arose in celestial mechanics, in the celebrated *non integrable* 3-body problem, involving the Sun, a big planet and a smaller one, of respective masses m_1, $m_2 < m_1$ and $m_3 \ll m_2$. The straightforward approach would be to consider the presence of the small planet as a small perturbation of the *integrable* 2-body problem for the masses m_1 and m_2. But when one tries to determine the solution as a series in powers of the mass ratio $\epsilon = m_3/m_2$, it appears unbounded terms, the so-called *secular terms*, increasing without bounds as fast as t, hence of ill-defined order and impairing the very consistency of the perturbation approach at long times $t > 1/\epsilon$. Accordingly, the perturbation expansion is *not uniformly convergent in time*, preventing from using it to investigate asymptotics and determine the fate of the 3-body system: the influence of the small planet on the motion of the bigger one, although seemingly a weak perturbation, might ultimately modifies its trajectory around the Sun, at least in some resonant cases.

The origin of secular terms lies in a phenomenon of *resonance*, that is best explained on an example: the *Duffing oscillator* $\ddot{x} + x = -\epsilon x^3$ with $\epsilon \ll 1$. When looking for a solution in the form $x(t) = \sum \epsilon^n x_n(t)$, each component $x_n(t)$ has to be bounded in order to get a consistent perturbation expansion, in which the hierarchy of terms of different orders remains valid forever: $\epsilon x_{n+1}(t) \ll x_n(t)$. These components should satisfy the following sequence of equations:

$$\ddot{x}_0 + x_0 = 0, \quad \ddot{x}_1 + x_1 = -x_0^3, \quad \ldots \qquad \text{(linearized operator } Lx \equiv \ddot{x} + x) \tag{1}$$

It comes $x_0(t) = ae^{it} + cc$ from which follows a secular contribution $(3i/2)a|a|^2\ t\ e^{it}$ in $x_1(t)$. In general, solving perturbatively $\dot{z} = f(z, \epsilon)$ for an expansion $z(\epsilon, t) = \sum_n \epsilon^n z_n(t)$ yields a hierarchical sequence of equations of the form $\dot{z}_n = Lz_n + \varphi_n(z_0, z_1, \ldots, z_{n-1})$ for $n \geq 1$, where $L = Df(z_0, \epsilon = 0)$ comes from the linearization in z_0 of the unperturbed evolution law. A secular divergence arises in z_n as soon as φ_n contains an additive contribution which is an eigenvector of L (part of a mathematical result known as the Fredholm alternative). The appearance of secular terms reflects a singular feature of the dynamics: the fact that the limits as $\epsilon \to 0$ and $t \to \infty$ do not commute. As a rule, such non-interversion is associated with generalized secular divergences: the fast, short-term dynamics finally contributes to the slow, long-term behavior. This feature is a clue towards using multiple-scale method.

2.2 Technical principles

The first step is to perform rescalings leading to dimensionless variables and functions, which evidence a small control parameter ϵ, related to scale separation and providing a natural parameter for a perturbation approach. The basic principle of multiple-scale method is to introduce additional *independent* time variables $t_1, t_2, \ldots t_n$ such that the physical situation corresponds in this extended

time-variable space to the line

$$t_0 = t, \quad t_1 = \epsilon t, \quad t_2 = \epsilon^2 t, \ \ldots \quad \text{hence} \quad \frac{d}{dt} = \frac{\partial}{\partial t_0} + \epsilon \frac{\partial}{\partial t_1} + \epsilon^2 \frac{\partial}{\partial t_2} + \ldots \tag{2}$$

It thus amounts to a perturbation expansion of the time-derivative operator. This method can be traced back to the *Lindstedt-Poincaré technique*, where the time variable t is expanded according to $t = s(1 + \epsilon\omega_1 + \epsilon^2\omega_2 + \ldots)$ and the evolution described in terms of the new variable s and unknown frequencies $(\omega_i)_{i\geq 1}$ to be determined self-consistently [Nayfeh 1973]. By contrast, the multiple-scale approach puts on a par $t_0 = t$ and the additional variables $(t_i)_{i\geq 1}$. The perturbation approach is then carried out as usual, plugging the above expression (eq. 2) of d/dt and the expansion $z(\epsilon, t) = \sum_{n\geq 0} \epsilon^n \ z_n(t_0, t_1, t_2, \ldots)$ into the evolution equation and identifying term-wise the coefficients of the successive powers of ϵ. The additional freedom thus introduced when considering $(t_i)_{i\geq 0}$ as independent variables will be compensated in the course of the computation, by imposing "solubility conditions" ensuring the vanishing of secular terms and the consistency of the perturbation method. In particular, it is possible to freely choose boundary conditions outside the physical line $t_1 = \epsilon t_0, \ \ldots, \ t_n = \epsilon^n t_0$. The resulting set of equations contains exactly the same information as the original one, only expressed in a different way: by construction, terms depending, say, on t_0, describe a fast component with no emerging slow trends that would intermix with the t_1-dependence; fast variables contribute only to fast modes. At the end, one restricts to the physical line thus turning back to the single "real" variable t. The benefit of the method is to provide a *joint access to dependences at different scales*, now expressing as dependences onto the different time variables $t_0, \ t_1 \ \ldots, \ t_n$. One introduces as many new variables as necessary to circumvent secular divergences. We have implicitly supposed above that the behavior at time scale $\Delta t = \mathcal{O}(1)$ corresponds to the fastest time scale of the evolution. If it were not the case, the rescaled time variables would be $t_0 = \epsilon^{n_0} t, \ t_1 = \epsilon^{n_0+1} t, \ \ldots$ if the fastest time scale is $\Delta t = \mathcal{O}(\epsilon^{n_0})$. More general time derivative expansion, associated with rescaled variables $t_n = \epsilon^{\alpha_n} t$ might be considered to better account for the hierarchy of characteristic time scales of the dynamics.

2.3 Multiple-scale method: abstract examples

Let us first consider the simplest possible example $\dot{x} = a(1 + \epsilon)x$, for which the exact solution is trivially known, allowing to appreciate the validity of the multi-scale approach compared to the straightforward perturbation expansion. In the latter case, one looks for a solution $x(t) = x_0(t) + \epsilon x_1(t) + \mathcal{O}(\epsilon^2)$ and identifies term-wise the powers of ϵ. At order 0, $\dot{x_0} = a x_0$ yields $x_0(t) = c_0 e^{at}$. At order 1, $\dot{x_1} - a x_1 = x_0(t)$ leads to a secular divergence: $x_1(t) = c_1^0 e^{at} + c_0 t e^{at}$. Carrying on the perturbation analysis yields the following expansion

$$x(t) = c \, e^{at} \, (1 + \epsilon t + \epsilon^2 t^2/2 + \ldots) \tag{3}$$

which is not uniformly convergent: for $t = \mathcal{O}(1/\epsilon)$, all terms are of the same magnitude (and for $t \geq 1/\epsilon$ the expansion no longer converges). *The method is only relevant at short times* $t \ll 1/\epsilon$. In the multiple-scale approach, one introduces two rescaled variables $t_0 = t$ et $t_1 = \epsilon t$ and looks for a solution of the form $x(t) \equiv x_0(t_0, t_1, \ldots) + \epsilon x_1(t_0, t_1, \ldots) + \mathcal{O}(\epsilon^2)$. At order 0, $\partial_{t_0} x_0 = a x_0$ yields $x_0(t_0, t_1, \ldots) = c_0(t_1, \ldots) e^{at_0}$. At order 1, we get $\partial_{t_0} x_1 + \partial_{t_1} x_0 = x_0 + a x_1$. The solubility condition writes $a c_0 - \partial_{t_1} c_0 = 0$, which allows to avoid secular divergence and suppresses the artificial freedom introduced with the additional time variable t_1, yielding $c_0 = c e^{at_1}$. The equation $(\partial_{t_0} - a) x_1 = 0$ is here superfluous, but in less simple situations, it remains at this stage a nontrivial equation for x_1. One thus directly gets the solution, valid at all times:

$$x(t) = c \, e^{at_1} \, e^{at_0} = c \, e^{a(1+\epsilon)t} \tag{4}$$

As a rule in singular perturbation method, the difficulty here originates in the non-commuting limits $\epsilon \to 0$ and $t \to \infty$; indeed, denoting $y_\epsilon(t) = x_\epsilon(t)e^{-at}$, one has $\lim_{t\to\infty}\ \lim_{\epsilon\to 0^+} y_\epsilon(t) = c$ whereas $\lim_{\epsilon\to 0^+}\ \lim_{t\to\infty} y_\epsilon(t) = \infty$.

Other training examples are the weakly damped linear oscillator $\ddot{x} + x = -2\epsilon\dot{x}$, solved with multiple scales $t_0 = t$, $t_1 = \epsilon t$, $t_1 = \epsilon^2 t$, or with the more specific variables $\theta = \sqrt{1-\epsilon^2}t$, $\tau = \epsilon t$; the Duffing oscillator $\ddot{x} + x = -\epsilon x^3$ introduced above, whose multiple-scale resolution requires three variables $t_0 = t$, $t_1 = \epsilon t$, $t_1 = \epsilon^2 t$; the Van der Pol oscillator $\ddot{x} + x = \epsilon(1-x^2)\dot{x}$.

2.4 An illustration: classical Lorentz electron gas in a weak field

As a less abstract hence more convincing illustration of the strength of multiple-scale method, let us consider the dynamics of a classical Lorentz electron gas acted upon an external electric field (associated acceleration **a**). This model considers the electrons as charged hard-spheres whose motion results from the superimposition of a driven classical motion in the field and elastic collision on immobile scatterers (the atoms). It is implemented within a kinetic-theoretic framework, based upon a Boltzmann-like equation for the electron velocity distribution:

$$\left(\frac{\partial}{\partial t} + \mathbf{a}.\frac{\partial}{\partial \mathbf{v}}\right) f(\mathbf{v},t) = -\frac{v}{\lambda}\mathcal{Q}f(\mathbf{v},t) \tag{5}$$

where $v = |\mathbf{v}|$ and λ is the mean free path of the electrons. $\mathcal{Q}f = f - f_{sph}$ is a projector accounting for the effect of collisions through the deviation of the distribution f from spherical symmetry, namely through the discrepancy between f and its isotropic counterpart $f_{sph}(v) = (1/4\pi)\int f(\mathbf{v},t)d\hat{v}$ obtained as an average over the velocity directions $\hat{v}$. The relevant small parameter is $\epsilon = ma\lambda/kT$, measuring the ratio of the work $ma\lambda$ done by the field over the mean free path to the thermal energy kT in the initial state. $\epsilon \ll 1$ ensures the separation of the characteristic time scales of the two mechanisms experienced by an electron: the thermal motion and the field-induced deterministic motion. Denoting $v_{th} = \sqrt{kT/m}$ the thermal velocity of the electrons, we have indeed $\epsilon = (t_{th}/t_{acc})^2$ where $t_{th} = \lambda v_{th}$ is the mean time between two successive collisions with the scatterers and $t_{acc} = \sqrt{\lambda/a}$ is the acceleration time required for the field to move the electron over the mean free path λ starting from rest. The result of the plain weak-field expansion is to evidence is own failure: it shows that the perturbation is singular insofar as the asymptotic state will be fully dominated by the field, with no memory of the initial temperature. Multiple-scale method is here implemented with respect to the time variable, introducing new independent variables $(\tau_i)_{i\geq 0}$ such that the physical situation corresponds to the line:

$$\tau_0 = tv_{th}/\lambda, \quad \tau_1 = \epsilon\tau_0, \quad \tau_2 = \epsilon^2\tau_0 \quad \dots \quad \tau_n = \epsilon^n\tau_0 \quad \dots \quad (\epsilon = ma\lambda/kT) \tag{6}$$

The time-derivative expansion (2) is supplemented with an expansion of the velocity distribution:

$$f(\mathbf{v},t) = \sum_{i\geq 0} \epsilon^i\ F^{(i)}(\mathbf{v},\tau_0,\tau_1,\dots,\tau_n,\dots) \tag{7}$$

The procedure is conducted as exposed in the general case. Identifying term-wise the coefficients of the expansion yields a hierarchy of equations for the $(F^{(i)})_{i\geq 1}$, each supplemented with a solubility condition preventing from the appearance of secular divergences. A detailed presentation can be found in [Piasecki 1993]. The benefit of the multiple-scale method is to yield *jointly* the different stages of the gas evolution, starting from thermal equilibrium and switching on the field at $t = 0$:
– at times $\tau = \mathcal{O}(1)$, an initial transient with a drift velocity $\langle v_z\rangle(t) = at - C_1 at^2 v_{th}/\lambda + \cdots$ in the direction of the applied field (denoting C_1 some numerical constant);
– at times $\tau = \mathcal{O}(1/\epsilon)$, a linear-response regime with a steady drift velocity $\langle v_z\rangle \sim a\lambda/v_{th}$;

– at times $\tau = \mathcal{O}(1/\epsilon^2)$, a long-time field-dominated heating of the gas, where the velocity distribution is no longer Maxwellian and the kinetic energy of the electrons grows without bounds as $t^{2/3}$, whereas the drift velocity slowly vanishes asymptotically: $\langle v_z \rangle \sim (\lambda^2 a/t)^{1/3}$.

2.5 Domains of application of the multiple-scale method

Multiple-scale method has been first developed in nonlinear mechanics. It is fruitful and even required in any instance where plain perturbation expansion is not uniformly convergent, more generally when it is necessary to account jointly for variations at different time scales: resonant wave interactions, e.g. in plasmas, or in case of oscillations with slowly varying coefficients. Multiple-time-scale method has been applied around 1960 to get kinetic equations (closed equations for the one-particle distribution) from molecular dynamics (Liouville equation) for dilute gases, plasmas, or to establish a microscopic theory of Brownian motion from molecular dynamics of a hard-sphere system (§ 7.2). In the same spirit, it allows to relate constructively different mesoscopic descriptions e.g., in the case of Brownian motion, to relate the Kramers equation for the distribution $P(\mathbf{r}, \mathbf{v}, t)$ to the Smoluchowski equation for $P(\mathbf{r}, t)$ (§ 7.3). Other examples are the determination of transport coefficients (friction, viscosity) from kinetic description, or at macroscopic scale, the determination of eddy viscosity and eddy diffusivity (§ 7.10). A last domain of application concerns systems where relaxation processes at different scales superimpose, requiring to handle jointly different time dependences. Multiple-scale method then displays the physics of the relaxation process and its associated hierarchical structure (see e.g. the application to the adiabatic piston problem, in this Encyclopedia [Gruber and Lesne 2006], also briefly sketched in § 5.3).

3 A brief overview of multi-scale approaches

3.1 Different scales and regimes

Common to all multi-scale approaches is the *focus on the very existence of different scales*, exploited through the use of rescaled variables which makes explicit the presence of a small parameter ϵ controlling the dynamics, responsible of the existence of different time scales and related to the scale separation. Technically, the first, very simple but essential step is to replace the variables, fields and parameters by their dimensionless counterparts. So doing, small parameters reflecting scale separation (in time, space, energies, amplitudes ...) will naturally appear. Although it is thus possible to estimate the order of the different terms, it is to be underlined that *it gives no clue on their actual contribution to the long-term behavior*: in singular situations, precisely those where multi-scale approaches have to be developed, small terms can have a noticeable influence at all scales. As illustrated in the following sections, different rescalings of variables and functions allow to discriminate features at different scales and to capture different regimes. More specifically, the techniques to manage with the joint contributions of several regimes at different time scales depend on the way these regimes intermix. They can be:

— either *superimposed* regimes, when fast and slow dependences intermingle in the evolution of the same variable. It is the frame of multiple-scale analysis. The solution writes typically $x(t, \epsilon t, \epsilon^2 t, \ldots)$.

— either *coexisting* regimes, namely a coexistence of fast and slow evolutions. One might focus either on the fast evolution and use a quasi-static approximation (or parametric approximation) for the slow evolution, either on the slow evolution and use a quasi-stationary approximation or an

averaging of the fast evolution. The solution writes typically $[x_{fast}(t),\, x_{slow}(\epsilon t)]$ (or $[x_{fast}(\tau/\epsilon),\, x_{slow}(\tau)]$ if the observation takes place at long time scales, with a relevant time variable $\tau = \epsilon t$).

— either *successive* regimes, when initial conditions, bulk behavior and asymptotics are not of the same order with respect to ϵ; this is a boundary-layer-like issue, and the solution writes typically $x_{layer}(t/\epsilon)$ for $0 \leq t \leq t_0$, then $x_{bulk}(t)$ for $t \geq t_0$, with $t_0 = \mathcal{O}(1)$.

Applications are innumerable; the most typical and investigated ones are the climate (from hours for the observed weather to thousand years for eras), population dynamics, coasts and sand dunes (from grains to country scales), protein folding (the vibration of covalent bonds occurs at scale of femtoseconds while the whole folding may require up to a few seconds), or trading markets (from seconds to years). Let us finally give two typical examples for the parameter ϵ:

– *the weak-damping and high-friction limits*, best explained on an example. The damped oscillator $m\ddot{x} + \gamma\dot{x} + V'(x) = 0$ appears as an *Hamiltonian dynamics* $m\ddot{x} + V'(x) = 0$ as soon as the damping can be neglected, when the characteristic time $\theta = [m/V''(0)]^{1/2}$ of the undamped oscillator is far smaller than the damping time $\tau = m/\gamma$. The weak-damping limit is thus defined as $\epsilon \to 0$ where $\epsilon = \theta/\tau = [\gamma^2/mV''(0)]^{1/2}$. It leads to a singular behavior when investigating the asymptotics, as in the Duffing oscillator and weakly damped oscillator mentioned in § 2.3. On the contrary, the evolution appears as a *dissipative gradient dynamics* $\dot{x} = -V'(x)/\gamma = 0$ as soon as $\tau \ll \theta$. This leads to the high-friction limit: $\tau/\theta = [mV''(0)/\gamma^2]^{1/2} \to 0$. This example somehow reconciles conservative and dissipative dynamics, showing that they might coexist in the same system.

– *the hydrodynamic limit* involved in the derivation of hydrodynamics equations (namely incompressible Navier-Stokes equations) from kinetic Boltzmann equation. It writes $\epsilon = \lambda/L \to 0$ where ϵ is the so-called *Knudsen number*, defined as the ratio of the mean free path λ (the average distance traveled by a fluid molecule between two successive collisions) to a characteristic spatial scale L of the system (e.g. the size of an obstacle).

3.2 Bridging the scales: mean-field, singular and scaling approaches

The aim of multi-scale approaches is to bridge different scales, either through the determination of the large-scale behavior of the solution, either by establishing a constructive relation between the initial model and an effective model at higher scale. We have yet mentioned in introduction a first classification multi-scale systems and associated approaches: they might exhibit either (i) scale decoupling; (ii) some singularity in the relation between the different scales; (iii) scale invariance.

Mean-field approaches

In case of scale decoupling, mean-field approaches apply. Let us briefly recall, within its usual spatial formulation, that a mean-field approach amounts to identify the local environment, which is a priori fluctuating and spatially inhomogeneous (e.g. the local magnetic field generated by neighboring spins in a spin lattice model) with the average one, expressed as a function of the average order parameter (spatial average or equivalently a statistical average in the limit as the system size tends to infinity). Mean-field approaches can be implemented either in time (*averaging*), in real space (*homogenization, coarse-graining*), or in phase space (*aggregation and projection techniques*).

In the present context, the best example of a mean-field approach is provided by *homogenization* procedures. They can be traced back to the method of Lagrange to solve the three-body problem. The issue is to describe the motion of a light body B_2 experiencing the gravitational attraction of the Sun and a heavier body B_1. The mass of B_2 is supposed to be enough small to neglect its

influence on the Sun and B_1 (the so-called restricted three-body problem); B_1 will thus obey the Keplerian laws of motion. The method of Lagrange applies when B_2 is far more distant from the Sun than B_1 ($r_2 \gg r_1$) which implies (due to the third law of Kepler: $\omega^2 r^3 = cte$) that the angular velocity ω_1 of B_1 is far larger than ω_2: the large body B_1 moves faster than B_2 around the Sun. In first approximation, Lagrange replaced the rapidly oscillating influence of B_1 on the motion of B_2 by the influence of a constant distribution of mass, obtained by spreading the mass m_1 of B_1 all over its orbit. The Gauss theorem thus states that this influence can be accounted for by simply adding the total mass of this distribution to the mass of the Sun. The stability of the system would follow: B_2 will remain trapped in the neighborhood of the pair composed with the Sun and B_1.

Singular perturbations

A typical instance of singular multi-scale behavior is associated with asymptotic expansions

$$x(t) = \sum_{r=0}^{n-1} \epsilon^r x_r + R_n(\epsilon, t) \tag{8}$$

which are *not* convergent: $\lim_{n\to\infty} R_n(\epsilon, t) \neq 0$ at ϵ fixed, but $\lim_{\epsilon\to 0} \epsilon^{-n} R_n(\epsilon, t) = 0$ at fixed n and t. Asymptotic expansions are ubiquitous in multi-scale approaches: the coexistence of different time scales, superimposed and non trivially coupled to get rise to the observed phenomenon, prevents from obtaining uniformly convergent perturbative expansions; it is only in this latter regular case that the above-mentioned mean-field approaches and homogenization techniques apply.

Scale invariance, scaling theories and renormalization

Self-similarity and associated criticality prevent from scale decoupling, but allows to develop scaling theories and renormalization methods. By contrast to scale-separation arguments, the guiding principle is now to focus on the links relating one scale to the other ones (scaling transformations, renormalization transformations). The problem complexity is thus reduced in a somehow "transverse way", by retaining only scale-invariant features. We shall expose in § 6 further links between multi-scale approaches and renormalization methods, beyond the restricted scope of scale-invariant systems: in many instances, renormalization can be seen as an iterated multi-scale approach.

3.3 Scaling limits

Let us mention a specific instance of multi-scale approach, that associated with scaling limits. Scaling limit refers to a joint limiting procedure, in which several independent variables *jointly* converge towards given limits, with *prescribed relative behaviors*; this latter condition is a key point in the frequent case when the different limits do not commute, and we shall see in § 6 that it is an essential ingredient of renormalization methods. Let us cite two acknowledged examples:

– *the thermodynamic limit* for a system of N particles in a volume V; it amounts to let $N \to \infty$, $V \to \infty$ while $N/V = n = cte$ (constant average number density). It is a prerequisite to derive standard thermodynamic behavior from the statistical-mechanical description; it supports the use of asymptotic results given by the law of large numbers and the central limit theorem provided the correlations between the particles remain short-range.

– *the Boltzmann-Grad limit* for a system of n hard spheres of radius ϵ per unit volume. In dimension d, it writes $\epsilon \to 0$, $n \to \infty$ (thus differing from the thermodynamic limit) while $n\epsilon^{d-1} = z$ remains constant. This limit is involved in kinetic theory as a limiting instance where the Boltzmann ansatz applies (identifying the two-particle distribution function with the product of the corresponding

one-particle distributions). Indeed, the occupied volume fraction $n\epsilon^d$ tends to 0 so that recollisions and ensuing long-term correlations can be neglected (rarefied gas). On the other hand, the mean free path of a particle remains finite, so that numerous collisions and associated molecular chaos further supports the Boltzmann decorrelation ansatz.

3.4 Stochastic multi-scale approaches

Multi-scale approaches are far less developed for stochastic processes. Let us mention the case of a Markov process. Scale separation reflects in a spectral gap in the transition matrix generating the dynamics. Identification of fast and slow modes is then straightforward: slow modes are associated with quasi-degenerated eigenvalues ($\lambda \approx 0$ in a time-continuous setting) whereas fast dynamics is associated with damped modes and negative eigenvalues ($\lambda < 0$, $|\lambda| \gg 1$) [Gaveau et al. 1999]. A basic difficulty in extending methods developed in a deterministic context is the fact that the reduction (or projection) of a Markov process is a priori no longer Markovian. Closure relations and approximations should be introduced to circumvent memory effects, e.g. supported by arguments of decorrelation and ensuing fast temporal self-averaging of the fast dynamics.

It is to note that the behavior upon rescaling of a stochastic process differs from the transformation of a deterministic evolution. The basic relation is the scaling upon a time rescaling $\theta = \epsilon t$ of the white noise involved in stochastic differential equations and defined from the Wiener process $W(t)$ through the relation $dW(t) = \eta(t)dt$. It follows from the definition $\widetilde{W}(\theta) = W(t)$ that $d\widetilde{W}(\theta) = \sqrt{\epsilon}\ dW(t)$. At this point, it is important to notice the difference with respect to the behavior of a plain deterministic function $\widetilde{f}(\theta) = f(t)$ for which $d\widetilde{f}(\theta) = \epsilon\ df(t)$. Using the fact that $\delta(t) = \epsilon\ \delta(\theta)$ and the definition $d\widetilde{W}(\theta) = \widetilde{\eta}(\theta)d\theta$, we obtain that $\widetilde{\eta}(\theta)$ is a white noise with respect to the rescaled time θ, i.e. a stationary Gaussian process defined by its first two moments

$$\langle\widetilde{\eta}(\theta)\rangle = 0 \qquad \langle\widetilde{\eta}(\theta)\widetilde{\eta}(\theta')\rangle = \delta(\theta - \theta') \tag{9}$$

4 Slow/fast variables

4.1 Slow/fast decomposition

Dynamics of systems made of many interacting elements, e.g. chemical reactions, or population dynamics, typically involves far too many degrees of freedom to be handled at the level of individual units, and requires a drastic reduction to make sense of it. A natural way of reduction is based upon the phenomenology, taking as relevant degrees of freedom those describing the slow evolution observed at macroscopic scales. Scale separation between microscopic and macroscopic worlds has to be turned into a constructive and quantitative argument to achieve this reduction.

Solving this typical multi-scale issue first requires *to identify and construct explicitely the slow variables*, e.g. collective variables obtained through aggregation or coarse-grainings. The second step is *to eliminate or rather integrate the fast dynamics* into a closed system of effective equations describing the large-scale evolution. The closure requirement generically involves an approximation, neglecting the remaining dynamic coupling between fast and slow variables. It is precisely here that scale-separation arguments and the very choice of the slow variables are crucial, ensuring that the influence of fast dynamics is essentially accounted for in its effective or average contribution to the slow dynamics; remaining fluctuating influences can be either neglected or included in a noise term, required to be fully determined as a function of the slow variable only (otherwise the whole

procedure would not be consistent nor useful). In the following subsections, we shall briefly present the main techniques allowing to achieve this program, considering the simple abstract system:

$$\left\{ \begin{array}{rcl} dX/dt & = & f(X,Y) \\ dY/dt & = & \epsilon\, g(X,Y) \end{array} \right. \qquad (\epsilon \ll 1) \qquad (10)$$

Although involving only two variables for simplicity, it exhibits the typical multi-scale structure: whereas X varies on scales $\mathcal{O}(1)$, Y appears as a slow variable of characteristic time-scale $\mathcal{O}(1/\epsilon)$.

4.2 Parametric approximation

The preliminary step of the reduction is to get some knowledge on the fast dynamics, at least to choose the proper multi-scale technique. A plain but nevertheless fruitful remark is that a parameter p can always be seen as a variable that does not evolve: $dp/dt = 0$ in a deterministic setting, or $W_{p\to q} = \delta(p-q)$ in a stochastic one (transition probability W). Conversely, a slow variable can be transiently treated as a mere parameter in the fast dynamics. Supported by time-scale separation, this *parametric approximation* (or *quasi-static* approximation) decouples the fast dynamics from the slow variable evolution, investigating the fast dynamics asymptotics ($t \to \infty$) while considering that the slow variable remains constant $Y(t) \equiv y$. We shall distinguish two cases: (i) the fast dynamics oscillates with a period $T \ll 1/\epsilon$ (§ 4.3, § 4.4) and (ii) the fast dynamics relaxes to a stable equilibrium point $X^*(y)$ slaved to the slow variable (§ 4.5, § 4.7, § 4.9).

4.3 Amplitude equations

An ubiquitous technique to account for *slowly modulated oscillations* has been introduced first by Fresnel for light propagation and optical phenomena. The basic idea is to take benefit from the scale separation between the fundamental oscillation (frequency ω, wavelength $\lambda = 2\pi/k$) and a superimposed slow variation of the wave amplitude

$$\mathcal{A}(\mathbf{r},t) = A(\mathbf{r},t)\, e^{i(\mathbf{k}.\mathbf{r}-\omega t)} \qquad \text{with} \quad K \equiv |\nabla A/A| \ll k \quad \text{and} \quad \Omega \equiv |\partial_t A/A| \ll \omega \qquad (11)$$

The evolution can be rewritten in terms of the slowly varying amplitude A; by construction it is ruled by terms involving the small parameter $\epsilon \sim K/k \sim \Omega/\omega \ll 1$, but the resulting equation is now devoid of small or large parameter. Such technique has been successfully applied and further developed e.g. in various situations involving *electromagnetic waves* (e.g. diffraction of Hertzian waves), in *plasma physics* (resonant interaction between electromagnetic waves and acoustic modes) and in *quantum mechanics*, to investigate the deformation of a wave packet in a potential.

4.4 Averaging

Let us discuss further, in a general setting, the case when the fast dynamics is an oscillation of period T (either linear modes as in § 4.3 or a stable limit cycle). It is a context where *averaging techniques* apply. We refer to the associated entry in this Encyclopedia [Neishtatdt 2006] and only mention here the main principle: to exploit scale separation and self-averaging property of the fast dynamics to replace $X(t)$ by an average value $X_{av}(t) = (1/T)\int_t^{T+t} X(s)ds$. The underlying idea is that averaging cancels out most of the fast variations so that $X_{av}(t)$ is now slowly varying. In case when the fast dynamics is influenced by the slow variable Y, its value is kept constant in

the averaging (parametric approximation, § 4.2). The resulting average behavior $X_{av}[Y(t),t]$ is reinjected in the evolution of the slow component, leading to a closed equation

$$\frac{dY}{dt} = \epsilon\, g\left(X_{av}[Y(t),t], Y\right) \qquad \text{or rather} \qquad \frac{d\widetilde{Y}}{d\tau} = g\left(\widetilde{X}_{av}[\widetilde{Y}(\tau),\tau], \widetilde{Y}\right) \tag{12}$$

in terms of the more relevant rescaled time variable $\tau = \epsilon t$ and $\widetilde{Y}(\tau) \equiv Y(t)$. Denoting $\bar{Y}(\tau)$ the solution of this approximate equation, the validity of the averaging procedure is assessed by theorems giving conditions ensuring that: $\lim_{\epsilon\to 0} \widetilde{Y}_\epsilon(\tau) = \bar{Y}(\tau)$. Note that such theorems (quite unusually) state the convergence, for a vanishing value of the perturbation parameter ϵ, of the exact solutions towards the approximate one (solution of the average equations).

To conclude, let us notice that one speaks of averaging in temporal context and homogenization in spatial or spatio-temporal contexts, when averaging is performed over space; as discussed in § 3.2, averaging and homogenization belongs to the general class of mean-field approximations.

4.5 Quasi-stationary approximation

Let us now consider the case when the fast dynamics converges at fixed Y towards a stable fixed point $X^*(Y)$. Focusing on the slow dynamics, the relevant time variable is $\tau = \epsilon t$, which turns the evolution (10) into

$$\left\{ \begin{array}{rcl} \epsilon\, dX/dt & = & f(X,Y) \\ dY/dt & = & g(X,Y) \end{array} \right. \tag{13}$$

(for the sake of simplicity, we use the same notation X for both $X(t)$ and $\widetilde{X}(\tau)$). It is solved in two steps, by noticing that at lowest order in ϵ, the fast dynamics reduces to the asymptotic regime $f(X,Y) = 0$, slaved to the slow variable Y. The corresponding stable state $X^*(Y)$ is then plugged into the slow dynamics to get a closed equation for $Y(\tau)$:

$$\frac{dY}{d\tau} = g[X^*(Y), Y] \equiv G(Y) \tag{14}$$

This achieves the desired dimensional reduction. It works equally well when X is a string of variables $X = (x_1, \ldots, x_N)$.

There is seemingly a paradox here, ubiquitous in many multi-scale approaches: in order to determine the evolution of the slow variable Y, it is considered a constant! The solution lies in scale separation: the trick is to consider the ensuing approximate decoupling as an exact one (what it would be in the limit $\epsilon \to 0$). In other words, the constancy of Y is considered over a time length which is long at the level of fast dynamics ($\Delta t \gg 1$), enough long for X to reach its equilibrium state $X^*(Y)$, but short at the macroscopic level ($\epsilon \Delta t = \Delta\tau \ll 1$). As in the so-called "quasi-static evolutions" encountered in thermodynamics, the large-scale evolution will be composed of a continued succession of local equilibrium states: at each time τ, X takes its instantaneous equilibrium value, slaved to $Y(\tau)$. One speaks equivalently here of *quasi-stationary approximation*, *quasi-steady-state approximation* or *adiabatic elimination of fast variables.*

4.6 Slow invariant manifolds

In the previous subsections, the decomposition between fast variables X and slow variables Y was given. But in practice, only the whole dynamics of the system is known and a main part of the issue is to find and construct explicitely the slow variables.

A geometrical viewpoint on the dynamics appears to be fruitful: if the system evolution is to be reducible to the evolution of a few degrees of freedom, it means that the flow essentially lives in a low-dimensional region of the phase space, that can be parametrized by these degrees of freedom up to some fuzziness of order $\mathcal{O}(\epsilon)$. Mathematical investigations have been conducted to assess this point, leading to the concept of *invariant slow manifold*: a manifold $\mathcal{M}$ of the phase space, invariant upon the dynamics and describing the slow dynamics once the system has reached it [Gorban et al. 2004]. Starting from an arbitrary point z_0, the trajectory first exhibits a fast transient bringing the system state close to $\mathcal{M}$, up to some tolerance of order $\mathcal{O}(\epsilon)$, then sticks to $\mathcal{M}$. Its evolution on $\mathcal{M}$ is ruled by a reduced dynamics, far slower than the fast relaxation to $\mathcal{M}$ as soon as the system actually exhibits a time-scale separation. This latter self-consistent assertion should be considered as a working hypothesis, to be validated by the explicit determination of $\mathcal{M}$ and associated reduced dynamics. This can be done numerically, by exploiting the presumed convergence property of any trajectory reaching $\mathcal{M}$ after some intrinsic transients. In other words, if the dynamics possesses a slow invariant manifold, an operational way to find $\mathcal{M}$ is to let the system evolves, starting from a sample of initial conditions, and to observe its stabilization on $\mathcal{M}$.

This framework obviously embeds the quasi-stationary approximation presented in § 4.5: in this case, the slow invariant manifold is $\mathcal{M} = \{z = (x, y), f(z) = 0\} = \{(x^*(y), y)\}$ and the dynamics restricted to $\mathcal{M}$ is the slow dynamics $dy/d\tau = G[y(\tau)]$, $x(\tau) = x^*[y(\tau)]$. The manifold is here invariant upon the approximate dynamics (for all t, $f[z(t)] = 0$ hence $z(t) \in \mathcal{M}$) but not upon the original one: some rigorous mathematical work has to be done to show that the actual dynamics keeps the trajectory in a proper neighborhood of $\mathcal{M}$ of width $\mathcal{O}(\epsilon)$. In other words, one has to control the discrepancy between the exact trajectory and the trajectory slaved on $\mathcal{M}$.

4.7 Central manifold

The notion of slow invariant manifold generalizes older results about *central manifolds*, exploited to reduce the dynamics *near a bifurcation point*. Let us consider a dynamical system $\dot{x} = f(x, \alpha)$ near a bifurcation point: in $\alpha = \alpha_c$, the fixed point x_0, stable for $\alpha < \alpha_c$, looses its stability. This reflects on the largest eigenvalue(s) of the stability matrix $Df(x_0, \alpha)$, namely $\lambda_1(\alpha) < 0$ for $\alpha < \alpha_c$, $\lambda_1(\alpha) > 0$ for $\alpha > \alpha_c$, and $\lambda_1(\alpha_c) = 0$. The small parameter is then $\epsilon = \lambda_1$. A main result was to show that near the bifurcation point, slow modes coincide with unstable directions and fast modes with stable directions [Haken 1983, 1996]. The decomposition in slow and fast variables is ruled by the *central manifold theorem*: the solutions can be expressed in terms of the amplitudes along the eigenvectors of the null-space of the dynamics at $\epsilon = 0$; these amplitudes appear as the relevant *order parameters* near the bifurcation. This is referred to as the *slaving principle*. Compared to the setting presented in § 4.6, the slow invariant manifold $\mathcal{M}$ is here given by the central manifold.

4.8 Projection techniques

The methods presented in the previous subsections to eliminate fast variables and construct a reduced slow dynamics can be unified into a common framework: *Mori-Zwanzig projection techniques*. The full state (x, y) of the system is projected onto the slow variable y and the functions $w(x, y)$ are projected onto their conditional expectation

$$\mathcal{P}w(y) \equiv \int w(x, y)\rho(x|y)dx \tag{15}$$

The core of the method lies in the choice of conditional distribution $\rho(x|y)$, for instance $\rho(x|y) = \delta(x - x^*(y))$ in case when there is an invariant manifold $x = x^*(y)$, or $\rho(x|y) = 1/2\pi$ in case of averaging over a rapidly varying phase x. We refer to [Givon et al. 2004] for a review.

4.9 Aggregation techniques and coarse-grainings

An intuitive guideline in the analysis of a multi-scale dynamics is that *collective variables or coherent states coincide with slow modes*. The rationale is that numerous fast fluctuations at the level of agent dynamics self-average, so that only a slow trend is perceptible at large scale. *Aggregation methods* have been developed in this spirit to build reduced models governing the slow dynamics [Auger 1989]. Nevertheless, in generic situations, aggregation does not led to closed equations for the collective variables and some level of approximation has to be introduced.

Let us here consider a system of N coupled degrees of freedom $[x_i(t)]_{I=1...N}$ (e.g. a system of N interacting agents) evolving deterministically according to a two-scale dynamics:

$$\epsilon \, \frac{dx_i}{dt} = f_i(x_1, \ldots, x_n) + \epsilon \, g_i(x_1, \ldots, x_n) \tag{16}$$

f describes a fast evolution due to the coupling between species and g_i a slow evolution due to internal mechanisms. A natural choice for the slow variable is $Y(x_1, \ldots, x_n) = \sum_i x_i$ but we shall write below the general case. The self-consistent requirement of the method is that this variable Y reflect a global and slow behavior [Auger and Bravo de la Parra 2000]. Considering t as a fast time variable, this condition amounts to require a quasi-static behavior for Y at this time scale. In other words, the consistency condition requires that it exists a manifold $\mathcal{F}_y$ such that

$$\sum_{i=1}^{N} \frac{\partial Y}{\partial x_i}(x_1, \ldots, x_N) \; f_i(x_1, \ldots, x_N) = 0 \quad \text{on} \quad \mathcal{F}_y = \{Y(x_1, \ldots, x_N) = y\} \tag{17}$$

We moreover assume that the fast dynamics on this manifold $\mathcal{F}_y$ leads to a stable equilibrium $(x_1^*(y), \ldots, x_N^*(y))$. We are then in position to describe the slow evolution of the manifold itself, i.e. the slow dynamics ruling the evolution of the aggregated variable y for ϵ enough small:

$$\frac{dy}{dt} = \sum_i \frac{\partial Y}{\partial x_i}[x_1^*(y), \ldots, x_N^*(y)] \; g_i[x_1^*(y), \ldots, x_N^*(y)] + \mathcal{O}(\epsilon) \tag{18}$$

Internal support of the procedure is to check the structural stability of this resulting aggregated dynamics. Compared to the quasi-stationary approximation and slaving principle presented in § 4.5, here the slow variable is not given independently but constructed as a function of the fast variables (aggregated variable). The same principles can also be implemented for discrete-time models.

Coarse-graining can be seen as the spatial analog of aggregation techniques developed in the phase space: the real space is split into cells considered as elementary units at macroscopic scale, and all the small-scale physics is averaged over each cell, yielding the apparent state of each unit (described by a few "coarse-grained" variables) and the effective interactions in between them.

Let us cite two hydrodynamic examples. *Eddy viscosity* refers to an effective viscosity involved in coarse-grained hydrodynamics equations; the contribution of small-scale turbulent structures is accounted for in an integrated way in this parameter, hence its name. It is typically lower than bare viscosity, even possibly reaching negative values at enough large Reynolds number, i.e. at enough low bare viscosities. *Cellular flows* are space-periodic flows, thus exhibiting a natural spatial scale: the coarse-graining amounts to an intrinsic homogenization over each cell of the flow.

Let us finally mention that coarse-grainings are involved in renormalization-group transformations once supplemented with the adequate rescalings (see § 6).

In conclusion, it is to note that all these various multi-scale approaches are closely related and can all be expressed as a specific projection technique in the extended phase space containing both fast and slow variables. For instance, aggregation techniques replacing the fast variables $(x_1, \ldots, x_n)$ by the slow collective variable $y = Y(x_1, \ldots, x_n)$ amounts to the projection technique involving the slow invariant manifold $\mathcal{M} = \{(x_1, \ldots, x_n, y) \mid y = Y(x_1, \ldots, x_n)\}$

4.10 Numerical aspects

In the community of applied mathematics, multi-scale methods refer specifically to numerical homogenization, involving *multigrid algorithms* as for instance multi-scale finite-element method, multigrid Monte Carlo, multigrid optimization or annealing. Basically, the idea of numerical homogenization is to avoid the numerical cost of using a mesh of size $h < \epsilon$ where ϵ is the scale of the smallest-scale features of the dynamics, and to use jointly:
— a fine mesh, to compute local quantities independently (hence with a parallelized program);
— a coarse mesh, to compute global behavior using effective parameters and homogenized quantities determined in the prior fine-mesh computation. We refer to [Gorban et al. 2004] for a review.

5 Boundary layers and matched expansions

5.1 Purposes and principles

Multi-scale approach to handle boundary layers has been introduced in 1905 by Prandtl in fluid mechanics for situations where the solution of hydrodynamics equations far from the boundaries ("bulk" solution) does not match the conditions at the surface of the walls or obstacles. This typically originates in the presence of a multiplicative small factor ϵ in front of the highest-order derivative; accordingly, the flow exhibits two different scales in space: a thin boundary layer of width controlled by ϵ and the bulk domain. The idea is to perform two different perturbation methods in the layer and in the bulk, involving a different rescaling in order to focus on and give the ruling place to either the boundary conditions, either the bulk dynamics (one also speaks of *inner* and *outer* expansions). Then these parallel perturbation expansions have to be bridged into a single global continuous solution. The matching principle is to identify the asymptotic behavior on the boundary side with the boundary condition of the bulk behavior [Nayfeh 1973]:

$$\lim_{r \to 0} X_{bulk}(r) = \lim_{\zeta \to \infty} X_{layer}(\zeta) \qquad \text{with} \qquad \zeta = r/\epsilon \tag{19}$$

Boundary layers of hydrodynamics have numerous analogs: *initial layers* in chemical kinetics, *skin layers* in electrodynamics and *edge layers* in solid state physics [Nayfeh 1973]. Adaptation of this technique is to be developed to determine the complete dynamics in the slow-invariant-manifold approach, matching the fast relaxation towards the manifold with the slow motion onto the manifold. Let us finally note that the matched expansion approach can benefit in each region of all the above-mentioned range of multi-scale techniques.

5.2 Time analog: implementation for initial layers

We shall here work out the time analog of a boundary-layer problem on the yet encountered abstract example (10), in the case when X rapidly evolves to a slaved equilibrium state $X^*(Y)$ but with initial conditions $Y(0) = y_0$ and $X(0) = x_0 \neq X^*(y_0)$. Obviously, the quasi-stationary approximation fails to described the initial regime and its applicability has to be reconsidered. The general principle of boundary-layer analysis, namely the recourse to two different perturbation approaches, is here implemented as follows:
— for the initial regime, one solves the fast dynamics with initial conditions $X(0) = x_0$ while keeping $Y(t) \equiv y_0$; this yields an approximate solution $[X_{layer}(t), Y_{layer}(t)]$, satisfying the initial conditions and valid at short times, as long as Y has not evolved;
— at longer times, the relevant variable is the rescaled time $\tau = \epsilon t$ and the quasi-stationary approximation described in § 4.5 applies.
The consistency of the two perturbative approaches is ensured by the *matching condition*:

$$\lim_{\tau\to 0} X_{bulk}(\tau) = \lim_{t\to\infty} X_{layer}(t) \qquad \lim_{\tau\to 0} Y_{bulk}(\tau) = \lim_{t\to\infty} Y_{layer}(t) \equiv y_0 \tag{20}$$

This condition is actually satisfied since $X_{bulk}(\tau) \equiv X^*[Y_{bulk}(\tau)]$ hence $\lim_{\tau\to 0} X_{bulk}(\tau) = X^*(y_0)$ and, by definition of X^* (at fixed $Y(t) \equiv y_0$) , $\lim_{t\to\infty} X_{layer}(t) = X^*(y_0)$.

5.3 Some typical applications

Enzymatic catalysis

A matched singular perturbation approach is currently encountered in chemical systems, for instance in the derivation of the *Michaelis-Menten kinetics* for a single enzyme and the *Hille cooperative kinetics* for an allosteric enzyme [Murray 2002]. Denoting E the enzyme, S the substrate, ES the active complex and P the product, the single-enzyme catalytic transformation of S into P is described by the following scheme:

$$S + E \underset{k'}{\overset{k}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} P + E \qquad [S] \equiv s, \quad [E] \equiv e \quad [ES] \equiv c \tag{21}$$

where, as well known, the enzyme is released at the end. Introducing dimensionless quantities

$$\tilde{t} \equiv k e_0 t, \quad \tilde{s} \equiv \frac{s}{s_0}, \quad \tilde{c} \equiv \frac{c}{e_0}, \quad K_m \equiv \frac{k' + k_{cat}}{k}, \quad \widetilde{K}_m = \frac{K_m}{s_0}, \quad \lambda \equiv \frac{k_{cat}}{k s_0}, \quad \epsilon = \frac{e_0}{s_0} \tag{22}$$

the corresponding chemical kinetic equations write

$$\frac{d\tilde{s}}{d\tilde{t}} = -\tilde{s} + \tilde{c}(\tilde{s} + \widetilde{K}_m - \lambda) \qquad \epsilon\, \frac{\tilde{c}}{d\tilde{t}} = g(\tilde{s}, \tilde{c}) \equiv \tilde{s} - \tilde{c}(\tilde{s} + \widetilde{K}_m) \tag{23}$$

Noticing that $\epsilon \ll 1$ (the enzyme is present in infinitesimal quantities compared to the substrate), a quasi-stationary approximation applies for the variable $\tilde{c}$: it means that the intermediary species ES rapidly reaches a local equilibrium state $\tilde{c} = \tilde{c}^*(\tilde{s})$. This yields the substrate evolution

$$\frac{d\tilde{s}}{d\tilde{t}} = \frac{\lambda \tilde{s}}{\tilde{s} + \widetilde{K}_m} \tag{24}$$

The initial condition is set only on the substrate: $s(0) = s_0$, i.e. $\tilde{s}(0) = 1$. It yields the well-known expression of the velocity $V \equiv (ds/dt)_{|t=0}$ as a function of the initial substrate concentration: $V(s_0) = e_0 k_{cat} s_0/(s_0 + K_m)$ (with a maximal value $V_{max} = e_0 k_{cat}$). The quasi-stationary value for the complex (dimensionless) concentration $\tilde{c}^*(\tilde{s} = 1) = 1/(1 + \widetilde{K}_m)$ at $t = 0$ obviously differs from the actual initial condition $\tilde{c}(0) = 0$: it is besides quite foreseeable that the transients leading the

complex ES to its stationary value cannot be described using a quasi-stationary approximation. At short times, the relevant time variable is the fast rescaled time $\theta = \tilde{t}/\epsilon$, leading to the equation describing the initial regime when supplemented with the actual initial condition $\tilde{c}(0) = 0$, $\tilde{s}(0) = 1$. The analysis is straightforwardly carried over, exactly as in the general abstract case, with a matching condition $\lim_{\theta\to\infty} \tilde{c}(\theta) = \tilde{c}(t = 0) = 1/(1 + \widetilde{K}_m)$.

Kinetic theory

Time matched expansions have been developed in kinetic theory, for instance to describe the fate of a tagged particle within a gas. In a first, short stage (*kinetic stage*) following the injection of the particle in the thermally equilibrated gas, the velocity distribution of the particle rapidly evolves due to collisions with gas molecules and associated momentum transfer. This stage lasts a few mean-free-times and it ends when the tagged-particle distribution is almost Maxwellian. Then in a second stage (*hydrodynamic stage*), the distribution slowly relaxes towards a spatially uniform distribution, ultimately equal to the equilibrium Maxwell-Boltzmann distribution; at each time, the velocity distribution is almost Maxwellian. The particle dynamics is described at the level of its distribution function by the Boltzmann equation, and the resolution (the so-called Chapman-Enskog method) is based on the above general principles [Dorfman 1999].

The adiabatic piston problem

A matched two-time-scale perturbation approach has been developed for the *adiabatic piston problem*: an isolated cylinder filled of an ideal gas (noninteracting light particles of mass m) is separated in two compartments by a moving piston, of mass M, adiabatic in the sense that it has no internal degrees of freedom and does not conduct heat when fixed. The small parameter is the mass ratio $\epsilon = 2m/(M + m)$. It quantifies the efficiency of energy transfer between the gas particles and the piston upon elastic collisions, and the strength of the indirect coupling of the two gas compartments through the collisions of their particles with one and the same piston. The matched perturbation approach gives access both to a fast deterministic relaxation towards mechanical equilibrium, at time scales $\mathcal{O}(1)$, with no heat transfer between the compartments, and a slow fluctuation-driven evolution towards thermal equilibrium, where the heat transfer is achieved by the collision-induced coupling between the gas and the piston fluctuating motion, thus occurring at time scales $\mathcal{O}(M/m)$ (see [Gruber et al. 2003] and the entry on the topic in this Encyclopedia, [Gruber and Lesne 2006]).

6 Renormalization: an iterated multi-scale approach

It is not the place to expose nor even summarize the implementation of renormalization techniques, for which I refer to the associated entries in this Encyclopedia. I will here only stress the natural relations between renormalization-group (RG) and multi-scale approaches: RG indeed shares many steps and guiding principles: joint rescalings, coarse-grainings and local averaging, effective parameters and effective terms, relevant and irrelevant contributions, with a focus on large-scale behavior. Moreover, far beyond the scope of critical phenomena study, RG has been extended into an *iterated multi-scale approach* allowing to determine in a systematic and constructive way the effective equation describing the universal large-scale features and asymptotics of a multi-scale system (see e.g. [Chen et al. 1996] [Oono 2000], [Mazzino et al. 2004]).

It is first to be underlined that different meanings are associated with the term "renormaliza-

tion", corresponding to very different statuses for the associated renormalization procedures.

A renormalized quantity can be plainly a *rescaled* quantity (normalized, dimensionless or put to the scale of the considered sample): here arises a first connection with multi-scale approaches, both involving rescalings as an essential preliminary step.

A renormalized quantity can be an *effective quantity* accounting in an integrated way of complicated underlying mechanisms (e.g. the renormalized mass of a body moving in a fluid, accounting for hydrodynamic effects); here arises another central notion of multi-scale approaches: effective parameters or effective equations (following e.g. from averaging or homogenization).

Renormalization is also a mathematical technique developed first in celestial mechanics then mainly in quantum electrodynamics to *regularize* divergent expansions and perturbation series. It might proceed by means of *resummation*; the idea, yet implemented by Rayleigh in 1917, is to sum up correlations and interactions into a redefinition of the parameters. It might either rely on the introduction of a *cutoff* in the space, time and energy scales, then accounting in an effective way of the host of contributions at smaller space-time scales $\Delta x \leq \Lambda$, $\Delta t \leq \theta$ (or equivalently larger momentum and frequency scales: $k \geq 2\pi/\Lambda$, $\omega \geq 2\pi/\theta$) so as to take advantage from physical cancellation of mathematical divergences. In any case, it turns the bare parameters of the original singular expansion into *renormalized parameters* and yields a renormalized regular expansion. Writing that the resulting large-scale behavior does not depend on the chosen cutoff (Λ, θ) yields *renormalization equations*, expressing quantitatively the very consistency of the procedure, ("renormalizability" of the expansion). Renormalization here provides alternative technical tools in instances treated above with multiple-scale method. Its main advantage is its *recursive structure*: introducing a sequence $(\Lambda_n, \theta_n)_n$ of cutoffs (what is called momentum-shell RG), the whole procedure can be iterated to integrate recursively the influence of small-scale features on the asymptotic behavior, allowing to handle situations exhibiting a hierarchy or even a continuum of scales.

Renormalization also refers to an *asymptotic analysis* allowing to classify critical behaviors, to determine quantitatively critical exponents and to handle associated divergences. Indeed, the above-mentionned multi-scale approaches fail near bifurcation points or critical points. In this case, scale separation is replaced by *scale invariance*. The key idea, underlying renormalization-group techniques is to shift the focus on the scaling procedure itself. The basic point is to construct a renormalization transformation, consisting in *joint coarse-grainings and rescalings*, thus relating two models *describing the same phenomenon at different scales* [Lesne 1998]; it puts forward their self-similar properties and associated scaling laws, while eliminating specific small-scale details having no consequences on the asymptotic, large-scale behavior. The set of renormalization transformations has a semi-group structure with respect to the rescaling factor (or plainly with respect to iteration) justifying to speak of the *renormalization group*. It generates a *flow in the space of models*, whose fixed points correspond either to trivial, either to critical situations according to their stability. It can be shown that the linear analysis of the renormalization transformation around a critical fixed point gives access to the critical exponents. This analysis moreover allows to split the space of models into *universality classes*, each associated to the basin of attraction of a critical fixed point. Let us emphasize that scale invariance leads to a deep change in the modeling and investigations, shifting from a physics focusing on the prediction of amplitudes to a "physics of the exponents", focusing on less specific, but more universal and above all, more intrisic features.

Far more generally, RG is associated with a qualitative change in the questioning, since the study takes place in a space of models. Generalized renormalization transformation can be designed

to extract not only self-similarity properties but any large-scale feature from a more microscopic model. In particular, RG can be specially designed to discriminate between essential and inessential terms in a model: the latter do not modify the asymptotics of the RG flow, meaning that they are of no consequence at large scales. In other words, generic properties of the renormalization flow in this space of models yields universal large-scale scaling properties. Renormalization-group is thus essentially a multi-scale approach, insofar as it only retains the relations between the different levels of descriptions, somehow ignoring the details at each one given scale. *It is actually designed to capture universal features of the multi-scale organization.*

7 Summary: the exemplary case of diffusion

7.1 Bridging the scales

Our aim in this section is to present at work the whole range of multi-scale approaches, allowing both to bridge models devised at different scales and to predict the large-scale features of the phenomenon they account for. We choose the context of diffusion, Brownian motion and transport phenomena, where such a bridge is essential and has been much investigated. Indeed, transport coefficients are defined through phenomenological equations; it is thus necessary to relate such macroscopic equations with smaller-scale theories, so as to get an expression of the coefficients in terms of the microscopic ingredients and to justify the validity of the phenomenological description.

The exposition, following increasing scales, will mark out the pathway from reversible molecular dynamics to macroscopic diffusion equations. We shall thus come across the multiple-scale analysis of the Liouville equation describing at microscopic scales a Brownian grain suspended in a thermal bath of water molecules (§ 7.2) leading to the mesoscopic Kramers equation for the grain distribution function $P(\mathbf{r}, \mathbf{v}, t)$. Involving higher but still mesoscopic scales, another multiple-scale analysis (§ 7.3) leads to the reduced Smoluchowski equation for its spatial distribution $P(\mathbf{r}, t)$. Random walks offer alternative mesoscopic models, involving effective diffusion coefficients in order to take into account underlying features like persistence length or other short-range correlations (§ 7.4). Scaling limits or more systematic renormalization methods in real space allow to bridge discrete random-walk models with continuous descriptions (§ 7.5). Another renormalization group, based on a path integral formulation in the frame of field theory, allows to handle the case of self-avoiding walks with infinite memory (§ 7.6). Homogenization is illustrated on the case of diffusion in a regular porous medium (§ 7.7) whereas diffusion processes in fractal substrates provide a counter-example, enough singular to exhibit anomalous scaling behavior (§ 7.8). The issue of reducing the dynamics of the diffusion process to a simpler effective one is encountered in many other macroscopic instances, among which we shall mention diffusion in a periodic medium, lending to space averaging (§ 7.9), and advection of a passive scalar field in a two-scale velocity field, where a multiple-scale analysis yields the effective diffusivity at large scale (§ 7.10). We shall give below further technical guidelines for constructing these stairs climbing from molecular up to large macroscopic scales, thus providing additional illustrations of the multi-scale approaches introduced in the previous sections on more general and abstract grounds.

7.2 Microscopic theory of Brownian motion

The first theoretical account of Brownian motion, namely the erratic movement of a micron-sized pollen grain suspended in a thermal bath, e.g. water, dates back to 1905 and the famous paper by Einstein. It took almost 60 years before a *microscopic* theory has been achieved [Lebowitz and Rubin 1963]. This theory has been further worked out using multiple-scale techniques [Cukier and Deutsch 1969]. The challenge is to start from the complete deterministic reversible dynamics of the system, described within a probabilistic framework by the Liouville equation $\partial p/\partial t = Lp$ for the distribution of probability p in the whole phase space (position and velocities of the grain, of mass M, and all water molecules, of mass $m \ll M$). The small parameter is the mass ratio $\epsilon = \sqrt{m/M}$ measuring the efficiency of the energy transfer upon collisions between the grain and the bath particles, assuming a binary interaction potential $U = \sum_i u(|\mathbf{r}_i - \mathbf{r}|)$. The Liouville operator is decomposed into $L = L_0 + \epsilon\, L_1$, and one introduces rescaled time variables $\tau_n = \epsilon^n\, t$, where $\tau_0 = t$ is the time scale of the fluid particle dynamics. Multiple-scale method is carried out according to the general scheme, leading to the so-called Kramers equation

$$\left(\frac{\partial}{\partial t} + \mathbf{v}.\frac{\partial}{\partial \mathbf{r}}\right) P(\mathbf{r},\mathbf{v},t) = \zeta\, \frac{\partial}{\partial \mathbf{v}} \left[\mathbf{v} + \frac{kT}{M}\frac{\partial}{\partial \mathbf{v}}\right] P(\mathbf{r},\mathbf{v},t) \tag{25}$$

where the friction coefficient is here explicitly given:

$$\zeta = \frac{1}{3MkT} \int_0^\infty \langle \mathbf{F}_t.\mathbf{F}_0 \rangle\, dt \quad \text{where} \quad \mathbf{F}_t = e^{iL_0 t}\mathbf{F}_0 \quad \text{and} \quad \mathbf{F}_0 = -\nabla_{\mathbf{r}} U \tag{26}$$

We refer to the original though very pedagogical paper [Cukier and Deutsch 1969] for a thorough exposition and discussion of this derivation.

7.3 Mesoscopic theory of Brownian motion

Multiple-scale method is also of relevance to determine the high-friction limit of the above Kramers equation. Standard perturbation technique with respect to the inverse $1/\zeta$ of the friction fails to describe the asymptotic regime: there is not enough freedom to fulfill all the solubility conditions required to avoid the appearance of secular divergences [Bocquet 1997]. By contrast, multiple-scale technique yields a uniform expansion of the evolution equation still valid at long times, thus allowing to bridge two mesoscopic levels of description, namely the Kramers equation and the Smoluchowski equation for the spatial density $\rho(\mathbf{r},t)$ of the Brownian particle

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) = \frac{1}{M\zeta}\, \frac{\partial}{\partial \mathbf{r}} \left(kT \frac{\partial}{\partial \mathbf{r}}\right) \rho(\mathbf{r},t) \tag{27}$$

Introducing dimensionless variables $\tau = t v_{th}/l$, $\mathbf{R} = \mathbf{r}/l$, $\mathbf{V} = \mathbf{v}/v_{th}$ where l is the size and $v_{th} = \sqrt{kT/M}$ the thermal velocity of the grain, the relevant small parameter appears to be the dimensionless inverse of the friction coefficient

$$\epsilon = \frac{v_{th}}{l\zeta} \quad \text{hence} \quad \epsilon \left(\frac{\partial}{\partial \tau} + \mathbf{V}.\frac{\partial}{\partial \mathbf{R}}\right) P(\mathbf{R},\mathbf{V},\tau) = \frac{\partial}{\partial \mathbf{V}} \left[\mathbf{V} + \frac{\partial}{\partial \mathbf{V}}\right] P(\mathbf{R},\mathbf{V},\tau) \tag{28}$$

If the friction is high (i.e. $\epsilon \ll 1$) the velocity relaxes very rapidly towards the equilibrium Maxwell distribution and it is then enough to describe the (slow) evolution of the spatial distribution $\rho(\mathbf{r},t)$. Nevertheless, the relaxation stage is essential and accordingly the ϵ-dependence is singular, as a rule when the small perturbation parameter multiplies the time derivative.

According to the general procedure exposed in §2, we introduce rescaled variables $\tau_0 = \tau$, $\tau_1 = \epsilon\tau$, $\tau_2 = \epsilon^2\tau$, ... considered as independent variables and look for a solution of the Kramers equation of the form $P = P^{(0)} + \epsilon P^{(1)} + \epsilon^2 P^{(2)} + \ldots$ where the arguments of all the components

$P^{(i)}$ are $(\mathbf{R}, \mathbf{V}, \tau_0, \tau_1, \tau_2, \ldots)$. Identifying term-wise the successive powers of ϵ yields a hierarchy of equations. At order 0, we obtain $P^{(0)} = \Phi(\mathbf{R}, \tau_0, \tau_1, \tau_2, \ldots)e^{-V^2/2}$. The following equations, for the $[P^{(i)}]_{i\geq 1}$, involve the linearized operator $\mathcal{L} = \partial_{\mathbf{V}}(\mathbf{V} + \partial_{\mathbf{V}})$. For each of them, it appears a solubility condition, requiring that none of the additive contributions appearing in the equation is an eigenvector of $\mathcal{L}$; involving the components $P^{(j)}$ with $j < i$, it prevents from the appearance of a secular divergence in $P^{(i)}$. At first order, the solubility condition is $\partial\Phi/\partial\tau_0 = 0$ thus determining the (trivial) τ_0-dependence of $P^{(0)}$. In a similar way, the solubility condition at order 2 allows to determine the τ_1-dependence of $P^{(0)}$. This achieves to bridge Kramers and Smoluchowski equations in the high-friction limit, when retaining only the first-order term in ϵ. We refer to [Bocquet 1997] for a pedagogical account of the derivation and discussion of its relation with the time-derivative expansion involved in the so-called Chapman-Enskog solution of Boltzmann equation.

7.4 Random-walk model and weakly correlated diffusion

Random walks are discrete-time mesoscopic models, accounting for the diffusing motion of a particle through the statistical properties of its successive steps, when observed at a given time scale τ. The basic model (*ideal random walk*) assumes isotropic, independent and identically distributed steps of variance a^2. Central limit theorem straightforwardly gives the time dependence of the mean-square-displacement $R^2(t) \equiv \langle|\mathbf{r}(t) - \mathbf{r}(0)|^2\rangle = a^2 t/\tau$, showing that the motion is a normal diffusion with diffusion coefficient $D = a^2/2d\tau$ in dimension d. It is to note (see also § 7.5) that D depends τ and a, but in a *joint manner*. Actually, the diffusion coefficient associated with a diffusive motion observed at scale a and modeled by a random walk on a lattice of parameter a can be written $D = \alpha a^2$ where the rate α depends on a (effective rate at spatial resolution a): here is involved a kind of renormalization to account in the rate $\alpha(a)$ of all microsteps backward and forward of length far smaller than a.

In case of short-range correlations between the successive steps (namely if $\sum_{-\infty}^{\infty} |C(t)| < \infty$ where $C(t)$ is the statistical correlation function between elementary steps separated by a time length t), direct computations support a time-average-like result: the asymptotic bahavior is still described by a normal diffusion law $R^2(t) \sim 2dD_{eff}t$ with $D_{eff} = D\sum_{-\infty}^{\infty} C(t)$. In case when $C(t) = e^{-t/\tau}$, $D_{eff} = D(1 + e^{-1/\tau})/(1 - e^{-1/\tau})$ hence $D_{eff} \approx 2\tau D$ if $\tau >> 1$.

7.5 Renormalization analysis in case of Markovian diffusion

Trying to bridge lattice random walks with a continuous description brings out the following difficulty: as the step size a goes to 0, one has obviously to decrease accordingly the duration τ, but choosing by which amount is not so obvious, since the walker velocity is ill-defined (it depends on the observation scale). Determination of the proper joint rescaling can be guessed from knowledge obtained by another mean about the system; rather, it can also be obtained in a systematic way thanks to RG methods. Let us explain the basic principle.

Let us denote $P_{a,\tau}(x, y, t)$ the transition probability ruling the random walk, namely the density of probability to jump from x to y in time t, where x, y are restricted to the lattice $(a\mathbf{Z})^d$ and time to $\tau\mathbf{N}$. The renormalization transformation $\Phi_{k,\alpha}$ should express the consequence on $P_{a,\tau}$ of a joint rescaling of space (by a factor of k) and time (by a factor of k^α). Taking into account the Markov character of the walks, we are thus led to define

$$[\Phi_{k,\alpha}P_{a,\tau}](x, y, t) \equiv k^d P_{a,\tau}(kx, ky, k^\alpha t) \qquad \text{in dimension } d \tag{29}$$

The proper value of α is to be determined self-sonsistently in order that the limit $\lim_{k\to\infty} \Phi_{k,\alpha} P_{a,\tau}$ exists (it is then a continuous transition probability $P^*_\alpha(x,y,t)$ defined on $\mathbf{R}^d \times \mathbf{R}^d \times \mathbf{R}$). The root-mean-square displacement $R(P,t) \equiv [\sum_{x,y} |x-y|^2 P(x,y,t)]^{1/2}$ is transformed according to

$$R(\Phi_{k,\alpha} P_{a,\tau}, t) = k^{-1} R(P_{a,\tau}, k^\alpha t) \tag{30}$$

Accordingly, it yields the diffusion law associated with the fixed point P^*_α

$$\text{for any } k, \quad R(P^*_\alpha, t) = k^{-1}\, R(P^*_\alpha, k^\alpha t) \quad \text{hence} \quad R(P^*_\alpha, t) \sim t^{1/\alpha} \tag{31}$$

It is anomalous except if $\alpha = 2$. In the case of ideal random walks, the proper exponent leading to a non trivial limit is $\alpha = 2$; this limit P^*_2 is the transition probability of a *Wiener process*

$$W_D(x,y,t) = [4\pi d D t]^{-d/2}\, e^{-(x-y)^2/4dDt} \quad \text{with} \quad D = a^2/2d\tau \tag{32}$$

This shows that *all ideal lattice random walks belong to the same universality class*, that of the Wiener process. This approach has been fruitfully applied to diffusion in disordered systems, the issue being to determine whether the disorder, accounted for as a noise term in the transition probabilities, modifies or not the normal diffusion law obtained in the unperturbed situation. Similar reasoning can also be implemented for self-similar anomalous diffusion processes, like *fractional Brownian motions* and *Levy flights* [Lesne 1998].

7.6 Renormalization analysis for self-avoiding walks

Let us only mention for the sake of completeness the renormalization techniques developed for determining the conformational statistics of linear polymer chains, whose 3-dimensional shape can be represented as the trajectory of a self-avoiding random walk. These techniques belong to the RG corpus developed in statistical mechanics for critical phase transitions, within a field-theoretic framework. A formal but exact analogy can actually be worked out between self-avoiding walks and a spin-lattice system with $n \to 0$ where n is the number of spin components.

The multi-scale nature of the system is here so marked that it should rather be qualified an *absence of characteristic scale*. In this respect, standard RG methods developed for critical phenomena lie at the very boundary of multi-scale approaches. Scale decoupling is replaced by scale invariance, which is somehow the conjugate situation: homogeneity in real space is replaced by homogeneity in the conjugate space (space of characteristic scales). Scale invariance here reflects in the self-similar property $R(N) \sim N^\nu$ relating the end-to-end distance R of the chain to the number N of elementary steps (the monomers), with an anomalous exponent ν (the Flory exponent $\nu \approx 3/5$ in dimension $d = 3$) originating from the infinite memory of the non-overlapping chain. We refer to [De Gennes 1984] for a thorough exposition of the concepts and techniques only alluded here.

7.7 Effective diffusion in a porous medium (homogenization)

Describing the diffusion in a porous medium appears as a formidable task at the pore level: it would require to account for all the boundary conditions at the border of the hollow domain $\mathcal{V} \in \mathcal{V}_0$ actually accessible to diffusion. When the pores have a finite characteristic size a, an homogenization approach can be developed at scales far larger than a. It allows to account for the slowing down of the motion due to obstacles in an effective diffusion coefficient (in plain words, the black and white medium made of matter and holes of size a appears as a grey homogenenous medium at larger scales). More specifically, a diffusing tracer of random trajectory $\mathbf{r}(t)$ experiences a varying coefficient $D[\mathbf{r}(t)]$ (it equals D inside the pores whereas it vanishes in the non-accessible

region $\mathcal{V}_0 - \mathcal{V}$). The idea is to replace this fluctuating realization of the transport coefficient by its spatial average (independent of the trajectory), in what concerns macroscopic properties:

$$D_{eff} = \int_{\mathcal{V}_0} D\, n_0(\mathbf{r}) d^d\mathbf{r} = \int_{\mathcal{V}} D[\mathbf{r}] d^d\mathbf{r} \qquad \text{(where } n_0(\mathbf{r}) = 1 \text{ iff } \mathbf{r} \in \mathcal{V}\text{)} \tag{33}$$

Rigorous mathematical theorems ensure that the large-scale motion can actually be described by a Fick law and associated plain diffusion equation [Bensoussan et al. 1978].

7.8 Anomalous diffusion in a fractal medium

The above homogenization for diffusion in a porous medium works well only if the pores have a finite characteristic size; by contrast, diffusion in a fractal substrate (e.g. a porous medium with pores of all sizes) generically leads to anomalous diffusion, associated with a time dependence of the mean-square displacement $R^2(t) \sim t^\gamma$ with $\gamma < 1$. In a fractal substrate, the existence of obstacles and pores of all sizes introduces spatial fluctuations at all scales and long-range correlations in the spatial dependence of D. This case corresponds to a critical situation and homogenization fails to give a relevant description of the macroscopic behavior, in the same way as mean-field methods fail to account for critical phase transitions. It reflects in the anomalous exponent $\gamma < 1$ of the diffusion law, that can be related to the fractal characteristics of the substrate ($\gamma = d_s/d_f$ where d_s is the spectral dimension and d_f the fractal dimension).

7.9 Effective diffusion in a periodic potential (averaging method)

In case of a periodic medium, where $D[\mathbf{r}(t)]$ oscillates with a small spatial period, an averaging procedure can be developed as in § 7.7, to determine an effective diffusion equation accounting for the large-scale motion. Explicit computations within a multiple-scale approach yield

$$D_{eff} = \frac{1}{\langle D \rangle} \tag{34}$$

where $\langle D \rangle$ denotes a space average over the elementary cell [Givon et al. 2004].

Let us rather detail the case of diffusion of a Brownian particle in a periodic potential U, with $U(x + L) = U(x)$ for any x (restricting to dimension 1 for simplicity), at equilibrium at temperature T. Let D be the coefficient of this particle in the absence of the potential. At large scales $dx \gg L$, the substrate appears to be spatially uniform. The influence of the periodic bias exerted by the potential on the diffusive motion (superimposition of a modulated deterministic drift) can be described in an average way. The result is a normal diffusion with a reduced effective diffusion coefficient

$$D_{eff}(U) = D \inf_{f \in \mathcal{C}^\infty(L\mathbf{S}_1)} \int_0^L |1 - f'(x)|^2 dm_U(x) \quad \text{with} \quad dm_U(x) = \frac{e^{-U(x)/kT} dx}{\int_0^L e^{-U(x')/kT} dx'} \tag{35}$$

where the infimum is taken over the set of smooth periodic functions of period L and the average involves the equilibrium distribution m_U of the particle in the potential landscape $U(.)$. So doing, one sees in particular here that no oriented motion can arise at equilibrium, even if U is asymmetric. The procedure extends to dimension d with only technical differences.

7.10 Effective diffusivity for a passively advected scalar

Still another fruitful implementation of multiple-scale method is encountered in the context of diffusion and transport phenomena, in the study of the advection by a given incompressible

velocity field $\mathbf{v}(\mathbf{r},t)$ of a passive scalar field $\theta(\mathbf{r},t)$, e.g. the density of small inert "tracer" particles advected by the fluid flow without modifying it back. We consider the case when the fluid motion can be decomposed into a large-scale, slowly varying component and a small-scale, rapidly varying fluctuation: $\mathbf{v}(\mathbf{r},t) = \mathbf{U}(\mathbf{r},t) + \lambda\mathbf{u}(\mathbf{r},t)$. The parameter λ controls the relative strength of these components. Another small parameter λ is involved in this problem: the ratio $\epsilon = l/L \ll 1$ of the typical length scales L and l of $\mathbf{U}$ and $\mathbf{u}$ respectively. The issue is here to bridge two macroscopic descriptions: the full hydrodynamic equation describing the evolution of the scalar field $\theta(\mathbf{r},t)$

$$\frac{\partial}{\partial t}\theta(\mathbf{r},t) + \mathbf{v}(\mathbf{r},t).\nabla\theta(\mathbf{r},t) = D\Delta\theta(\mathbf{r},t) \tag{36}$$

and a large-scale effective transport equation for an average scalar field $\theta_L(\mathbf{r},t)$

$$\frac{\partial}{\partial t}\theta_L(\mathbf{r},t) + \mathbf{U}(\mathbf{r},t).\nabla\theta_L(\mathbf{r},t) = \frac{\partial}{\partial r_i}\left[D_{ij}^{eff}\frac{\partial}{\partial r_j}(\mathbf{r},t)\theta_L(\mathbf{r},t)\right] \tag{37}$$

This procedure, amounting to account in an average way of the small-scale contributions to the complete hydrodynamic description, relies on a spatio-temporal generalization of the multiple-scale method: it involves rescaled space and time variables $\mathbf{X} = \epsilon\ \mathbf{x}$, $\tau = \epsilon\ t$, $T = \epsilon^2\ t$ The different characteristic scales of the velocity components are directly reflected in their arguments: $\mathbf{u}(\mathbf{x},t)$ and $\mathbf{U}(\mathbf{X},T)$. The passive scalar field now expresses $\theta(\mathbf{x},t,\mathbf{X},\tau,T)$ and it is expanded $\theta = \theta^0 + \epsilon\ \theta^1 + \epsilon^2\ \theta^2$ The standard multiple-scale procedure leads to introduce an auxiliary field χ

$$\partial_t\chi_j + [(\mathbf{u} + \lambda\mathbf{U}).\partial]\chi_j - D\ \partial^2\chi_j = -\ u_j \tag{38}$$

yielding the effective diffusivity tensor (where $\langle\ \rangle$ is a space average)

$$D_{ij}^E \equiv \frac{D_{ij}^{eff} - D_{ji}^{eff}}{2} = D\sum_p \langle\partial_p\chi_i\partial_p\chi_j\rangle \tag{39}$$

Advection enhances transport and eddy diffusivity will be larger than molecular diffusivity. In realistic cases, there is a continuum of scales $u = \sum_{n=0}^{N} u_n$, where u_n has a characteristic scale $l_n \sim 2^{-n}l_0$. Multiple-scale method is to be iterated into a renormalization-group analysis, achieving a recursive integration of the small andfast scales into D^E starting by the smallest and fastest ones.

8 Conclusions

Multi-scale approaches allow to predict large-scale behavior generated by a given model; even more, they offer constructive tools to bridge models at different scales for the same phenomenon. They provide systematic and mathematically well-controlled tools to turn faithful but intractable models into effective reduced ones, thus lying at the core of statistical mechanics, many-body dynamical systems and more generally all issues of the still-in-progress complex systems science. Indeed, in complex systems (that might be their very definition), levels are so inter-related that it is essential to investigate jointly all the scales from elementary units up to the whole system and its emergent properties; neither theoretical nor numerical approaches can alone consider all the levels together, showing the relevance if not the necessity of multi-scale approaches.

Basic preliminary issues are to determine the proper elementary level, the proper collective variables, and the relevant small parameters. Let us remark that the implementation of a multi-scale technique rapidly faces the fundamental issue of defining what is a macroscopic variable; it offers some clues, indicating that a macroscopic variables might be a phenomenological quantity observable at our scale, a slow mode, or collective variable.

Multi-scale approaches take benefit of the separation of scales involved in the different mechanisms at work in the considered phenomenon. The basic idea, seen above at work in various instances and different ways, is to somehow decouple the different scales and to solve several simpler single-scale problems. Any multi-scale implementation actually involves, at some stage and more or less explicitly, a limiting process in which the scale separation ratio $1/\epsilon$ tends to ∞: this limiting process has to be carefully controlled in order that the method can be applied to real situation. Finally, to be successful, multi-scale approaches should achieve a trade-off between
— *accuracy* (minimizing the loss of information involved in the reduction or projection technique)
— *efficiency and tractability* (this is e.g. one of the major success of hydrodynamics)
— *robustness* of the resulting reduced model (to be checked a posteriori)
— *flexibility* (extending to heterogeneous systems involving different components)
— *scope* (bridging many different levels in order to capture the whole hierarchical structure)

Let us conclude by emphasizing a much fruitful benefit of multi-scale approaches: they allow to investigate structural stability of a model, in particular to evidence relevant parameters and essential mechanisms controlling large-scale features. In this respect, they lead beyond the (necessarily restricted) scope of a specific model and give an explicit account of the observer biased view, related to its scale of observation. They hence contribute to capture a more complete and controlled understanding of the real physical systems.

Related entries in the Encyclopedia

Adiabatic piston (131)
Averaging methods (402)
Boltzmann equation (136)
Dynamics in disordered systems (65)
Interacting particle systems and hydrodynamics (128)
KdV and other modulation equation (371)
Macroscopic equations in equilibrium and non equilibrium statistical mechanics (138)
Renormalization (167)
Renormalization-group and second-order phase transition (123)
Stability problems in celestial mechanics (99)

Bibliography

Technical details and several applications of multi-scale perturbative expansions, in particular multiple time-scale method, with references to the original papers, can be found in [Nayfeh 1973]; see also [Bender and Orszag 1978]. Applications of multiple-scale method, fully worked out in a very pedagogical way, can be found in the [Cukier and Deutsch 1969], [Piasecki 1993] [Bocquet 1997] and [Mazzino et al. 2004]. An acknowledged reference on homogenization techniques and multi-scale analysis in periodic media is [Bensoussan et al. 1978]; see also the monographs of [Lochak and Meunier 1988] and [Berdichersky et al. 1999]. Two recent review papers on multi-scale approaches and reduction techniques are [Givon et al. 2004] and [Gorban et al. 2004]. Basic principles and technical aspects of scaling theories and renormalization-group approaches can be found in [Goldenfeld 1992] and [Lesne 1998].

References

Auger, P. (1989) *Dynamics and thermodynamics of hierarchically organized systems*, Pergamon Press, Oxford.

Auger, P. and Bravo de la Parra, R. (2000) Methods of aggregation of variables in population dynamics, *C.R. Acad. Sci. Paris, Life Sciences* **323**, 665–674.

Bender, C.M. and Orszag, S.A. (1978) *Advanced Mathematical Methods for Scientists and Engineers*, McGraw Hill, New York.

Bensoussan A., Lions, J.L. and Papanicolaou, G. (1978) *Asymptotic analysis for periodic structures*, North Holland, Amsterdam.

Berdichersky, V. Jikov, V., and Paparieolasu, G. (1999) *Homogenization*, World Scientific, Singapore.

Bocquet, L. (1997) High-friction limit of the Kramers equation: The multiple time-scale approach, *Am. J. Phys.* **65**, 140–144.

Chen, L.Y. Goldenfeld, N., and Oono, Y. (1996) Renormalization group and singular perturbations: Multiple scales, boundary layers, and reductive perturbation theory, *Phys. Rev. E* **54**, 376–394.

Cukier R.I. and Deutch, J.M. (1969) Microscopic theory of Brownian motion: The multiple-time-scale point of view, *Physical Review* **177**, 240–244.

De Gennes P.G. (1984) *Scaling concepts in polymer physics*, 2nd edition, Cornell Univ. Press, Ithaca.

Dorfman, J.R. (1999) *An introduction to chaos in non equilibrium statistical mechanics*, Cambridge University Press.

Gaveau, B., Lesne, A., and Schulman, L.S. (1999) Spectral signatures of hierarchical relaxation, *Physical Letters A* **258**, 222-228.

Givon, D., Kupferman, R., and Stuart, A. (2004) Extracting macroscopic dynamics: model problems and algorithms, *Nonlinearity* **17**, R55–R127.

Goldenfeld, N. (1992) *Lectures on phase transitions and the renormalization group*, Addison-Wesley, Reading MA.

Gorban, A., Karlin, I., and Zinovyev, A. (2004) Constructive methods of invariant manifolds for kinetic problems, *Physics Reports* **396**, 197-403.

Gruber, C. and Lesne, A. (2006) Adiabatic piston, *Encyclopedia of Mathematical Physics*, Academic Press.

Gruber, C. Pache, S., and Lesne, A. (2003) Two-time-scale relaxation towards thermal equilibrium of the enigmatic piston, *J. Stat. Phys.* **112**, 1199–1228.

Haken, H. (1983) *Synergetics: An introduction*, 3rd edition, Springer, Berlin.

Haken, H. (1996) Slaving principle revisited, *Physica D* **97**, 95–103.

Lebowitz, J.L. and Rubin, E. (1963) Dynamical study of Brownian motion, *Phys. Rev.* **131**, 2381–2396.

Lesne, A. (1998) *Renormalization methods*, Wiley, New York.

Lochak, P. and Meunier, C. (1988) *Multiphase averaging for classical systems*, Springer.

Mazzino, A., Musacchio, S. and Vulpiani, A. (2004) Multiple-scale analysis and renormalization for preasymptotic scalar transport, *Phys. Rev. E* **71**, 011113.

Murray, J.D. (2002) *Mathematical biology*, 3rd edition, Springer.

Nayfeh, A.H. (1973) *Perturbation Methods*, Wiley, New York.

Neishtadt, A. (2006) Averaging methods, *Encyclopedia of Mathematical Physics*, Academic Press.

Oono, Y. (2000) Renormalization and asymptotics, *Int. J. Mod. Phys. B* **14**, 1327–1361.

Piasecki, J. (1993) Time scales in the dynamics of the Lorentz electron gas, *Am. J. Phys.* **61**, 718–722.

Table des matières