1 Model Hierarchies by (formal) Asymptotic Expansions / Perturbation methods

1.1 Scaling ("Adimensionalization" = "Nondimensionalization") of equations - Dimensional analysis

Mathematical models/equations yield formal connections between quantities in physics, chemistry, biology,... that are represented by numbers, matrices etc. At first, these quantities bear "units", for example the unknown dependent variable $v(x,t)$ is a velocity measured in "kilometers per hour" (km/h), the independent variable x is a length measured in "meter" (in the "SI system"), the independent variable t, the time, has the unit "second", etc. In addition to the "variables", models usually contain "parameters" that are "arbitrary, but fixed".

The bracket notation $[x]$ denotes that we take the "physical unit" of a quantity, e.g. $[x] = m$ ("m like meter" as unit of lenght) or $[t] = s$ ("s like second").
For a dimensionless quantity we have $[q] = 1$; "q has unit 1".

Of course we have to make sure that all added terms in a model equation and its right and left hand side consistently bear the same unit (a trivial test if a model equation can be correct). More than that, in order to do mathematical analysis or numerical computations we have to "adimensionalize" the equation, i.e. to get rid of all ("physical") units by a procedure called "scaling".
Such a scaling is performed by dividing all variables by some reference quantity with the same unit, such that the new quantity "has unit 1".
In this procedure also all "parameters" automatically become dimensionless.

Note that such a scaling means much more than simply "chosing the appropriate units"; e.g. when "atomic" or "astronomic" units are chosen, like "Angström" ($= 10^{-10}$ m) or "lightyears" as unit of length.
Further, this kind of scaling makes a "dimensional analysis" possible that yields insight into the connection and importance of the (physical) quantities in the model. It is the basis for a "scaling analysis" which is a fundamental tool of applied mathematics for the motivation/derivation of "reduced", "simpler" models for complex problems, yielding "model hierarchies". In some situations it is possible to make this "formal asymptotics" rigorous by "asymptotic analysis", where the convergence of the solution of the exact problem to the solution of the reduced / approximate model(s) is proven in the sense of theorems.

We start with a seemingly simple example of the "scaling" :

Example 1. "The projectile problem with non-constant gravity"

We consider the movement of a body thrown or shot vertically upwards from the surface of the earth, and want to determine how long it takes until it falls back to the ground. We consider a situation where we throw it really high, such that the "weight" of the body starts to differ from the weight at the surface of the earth.

The body will move along a "ray" through the center of the earth, and we denote its distance from the earth surface at time t^* by $x^*(t^*)$.

We assume that only gravity is acting on the body, and making many model simplifications, in particular neglecting the "Coriolis force", the function x^* solves the initial value problem (resulting from Newton's second law):

$$\frac{d^2x^*}{dt^{*2}} = -\frac{gR^2}{(x^*+R)^2}\,, \tag{1}$$

$$x^*(0) = 0\,, \qquad \frac{dx^*}{dt^*}(0) = V\,, \tag{2}$$

where g is the gravitational acceleration at the surface of the earth, R is the radius of the earth, and V is the initial speed. Of course some effects have been neglected, most importantly drag = "air resistance".

We want to compute the "trajectory" $x^*(t^*)$ and, e.g. the time T^* such that $x^*(T^*) = 0$. Note that we use the notation $x^*(t^*)$ for the "unscaled" problem in order to free the notation $x(t)$ for the "scaled" problem on which we do the mathematics/numerics. Physicists and other applied scientists tend to use the same symbols for the scaled and the unscaled equations.

Nondimensionalization ("Adimensionalization") of mathematical model problems consists of two steps:

Step 1: ***Preparation*** - Produce a list of all variables and parameters of the problem together with their physical dimensions.

Using the "SI system" ("Système Internationale d'Unités": meter, kilogram, second etc), we get for problem (1)-(2)

Quantity	Dimension	Unit
x^*	Length	m
t^*	Time	s
g	Acceleration	m s^{-2}
R	Length	m
V	Velocity	m s^{-1}

Table 1: variable - parameter list for example 1

Step 2: ***Scaling*** **= "divide by reference quantities":** For every variable,
a) choose as a unit an *intrinsic* **reference quantity**, i.e. a combination of parameters of the problem with the same dimension.
b) Then **divide** the variables by the respective reference quantities.

In general, there are several ways to find reference quantities, one can construct new ones from the parameters contained in the problem and one is allowed also to add reference quantities.
If possible, intrinsic reference quantities should be chosen in a way that they the "expected sizes" of the variables, such that you can expect to compute with "moderate" numbers.

After carrying out this procedure, the problem contains only dimensionless quantities, in particular only *dimensionless parameters*, i.e. dimensionless "combinations" of the original parameters.

Example 1 (continued)**.** In the example of (1), (2) there is more than one way to define reference lengths and times from the parameters.

If we assume a 'relatively small' (relative to what?) initial velocity, we expect the maximal height to be "small" in comparison to the radius of the earth. So in the "easy/obvious" choice of R as reference length, the requirement of a matching order of magnitude is not satisfied, which might not be "good" (in which sense?).

Example 1 - Scaling 1: As a first choice (#1) of reference length and time we take the most obvious choice and use the only direct reference length R and the reference time R/V. Dividing the unscaled variables by these "units" yields the dimensionless quantities "position" y and "time" τ :

$$y = \frac{x^*}{R}\,, \quad \tau = \frac{t^*}{R/V} \tag{3}$$

and the problem for the scaled dependent variable $y(\tau)$ takes the form

$$\varepsilon \frac{d^2y}{d\tau^2} = -\frac{1}{(y+1)^2}\,, \tag{4}$$

$$y(0) = 0\,, \qquad \frac{d}{d\tau}y(0) = 1\,, \tag{5}$$

and contains the dimensionless parameter ε :

$$\varepsilon := \frac{V^2}{gR}\,. \tag{6}$$

Example 1 - Scaling 2: Our second choice (#2) is the same reference length R and a "more elaborate" reference time:

$$z = \frac{x^*}{R}\,, \quad \sigma = \frac{t^*}{\sqrt{R/g}} \tag{7}$$

Then the nondimensionalized problem for the scaled dependent variable $z(\sigma)$ is

$$\frac{d^2z}{d\tau_1^2} = -\frac{1}{(z+1)^2}\,, \tag{8}$$

$$z(0) = 0\,, \quad \frac{d}{d\sigma}z(0) = \sqrt{\varepsilon}\,, \tag{9}$$

where we obtain the same dimensionless parameter ε as in (6).

Example 1 - Scaling 3: Now we make an effort to construct a reference length "more elaborate" than R, namely V^2/g, which also avoids the expected magnitude mismatch, and use a third reference time V/g:

$$x = \frac{x^*}{V^2/g}\,, \quad t = \frac{t^*}{V/g}\,, \qquad \text{and hence} \quad x(t) := \frac{x^*(Vt/g)}{V^2/g}\,. \tag{10}$$

The problem (1),(2) for $x(t)$ in scaled form thus reads

$$x'' = -\frac{1}{(1+\varepsilon x)^2}\,, \tag{11}$$

$$x(0) = 0\,, \qquad x'(0) = 1\,, \tag{12}$$

with, again, the same dimensionless parameter ε defined in (6).

Now we can say what 'relatively small' initial velocity actually means, namely that V is small compared to $\sqrt{gR}$, i.e. $V \ll \sqrt{gR}$, or $\varepsilon \ll 1$.

Note that in each scaling the number of parameters is reduced from 3 to only one and same dimensionless parameter ε given by (6), hinting to an intrinsic property of the system.

The 3 initial value problems (4)-(5), (8)-(9) and (11)-(12) are equivalent for finite $\varepsilon > 0$; however if we consider the approximations when we set $\varepsilon = 0$ (the "reduced problems"), the results will differ very much.

For any of these scalings we can compute the (ε dependent) height as a function of time. To obtain the time T^* until the body falls back to the ground in the dimensional formulation, one has to solve $z(\sigma) = 0$, $y(\tau) = 0$ or $x(t) = 0$ to get the non-dimensional T, and then scale back to get the correct answer in dimensional units. Note that both the scaled T and the original T^* depend on ε.

A general investigation for carrying out Step 2 relies on the following general statement about dimensional variables and parameters, which can be seen as an "axiom of dimensional analysis", sometimes called Buckingham-"Theorem":

Buckingham-"Theorem"*: Mathematical models of physical (chemical, biological, ...) phenomena can be nondimensionalized. Dimensionless parameters and intrinsic reference quantities can be chosen as products of powers of the original parameters.*

Of course, this heuristic statement is not a "theorem" in the strict mathematical sense; under certain conditions it can be made more precise and proven. Note that in Step 2 above, it is not a priori obvious that the re-scaling to dimensionless variables results also in dimensionless parameters. According to this "theorem", this fact is true (for correct models).

Typically, there are certain relations between dimensions of variables and parameters, e.g. Velocity $= \frac{\text{Length}}{\text{Time}}$. These become clear when all dimensions are expressed in terms of a set of *fundamental dimensions*, e.g. the **SI-system** (the international system of units), consisting of a choice of **7 units for 7 fundamental dimensions**, i.e. **meter** (m) for *length,* **kilogram** (kg) for *mass,* **second** (s) for *time,* **Ampere** (A) for *electric current,* **Kelvin** (K) for *temperature,* **mole** (mol) for *amount of substance,* and **candela** (cd) for *luminous intensity.*

For models in mechanics, like in the example above, the 3 fundamental dimensions length, mass, and time are sufficient. We stay with this situation and abbreviate the dimensions length, mass, and time by L, M, and T, respectively. Suppose we have a "mechanics" problem with N parameters $\alpha_1, \ldots, \alpha_N$, with dimensions $[\alpha_1], \ldots, [\alpha_N]$. They can be written as

$$[\alpha_k] = \mathrm{L}^{l_k}\ \mathrm{M}^{m_k}\ \mathrm{T}^{t_k}\,, \tag{13}$$

with $(l_k, m_k, t_k) \in \mathbb{R}^3$, $k = 1, \ldots, N$. Thus the dimension of a quantity is characterized by a 3-vector, e.g. velocity by $(1, 0, -1)$. Suppose we are looking for an intrinsic reference quantity for a variable α, whose dimension is characterized by the 3-vector (l, m, t). As a consequence of the "Buckingham theorem" above, there exists $(a_1, \ldots, a_N) \in \mathbb{R}^N$ such that

$$[\alpha] = \left[\prod_{k=1}^{N} \alpha_k^{a_k}\right]\,,$$

which is equivalent to

$$\mathrm{L}^l\mathrm{M}^m\mathrm{T}^t = \prod_{k=1}^{N} \mathrm{L}^{a_k l_k}\mathrm{M}^{a_k m_k}\mathrm{T}^{a_k t_k} = \mathrm{L}^{\sum a_k l_k}\mathrm{M}^{\sum a_k m_k}\mathrm{T}^{\sum a_k t_k}\,, \tag{14}$$

and, thus, to the linear system

$$\sum_{k=1}^{N} a_k l_k = l\,, \qquad \sum_{k=1}^{N} a_k m_k = m\,, \qquad \sum_{k=1}^{N} a_k t_k = t\,, \tag{15}$$

for the exponent vector $(a_1, \ldots, a_N)$. This system is solvable thanks to the "Buckingham Theorem".

Dimensionless parameters correspond to elements of the "nullspace" of the coefficient matrix

$$\begin{pmatrix} l_1 & \ldots & l_N \\ m_1 & \ldots & m_N \\ t_1 & \ldots & t_N \end{pmatrix}. \tag{16}$$

Two linearly dependent elements of the nullspace correspond to two dimensionless parameters β and γ, where one can be expressed in terms of the other by the relation $\gamma = \beta^\lambda$, for some exponent λ. These considerations yield the following

Corollary to "Buckingham Theorem": *Independently of the precise way to adimensionalize the problem, the dimensionless parameters are the same (up to "linear dependence"), since the nullspace of the "coefficient matrix" (of the powers) - like in example (16) - is determined only by the coefficient matrix.*
The number N^* of independent dimensionless parameters is equal to the dimension of the nullspace. In our "mechanics" example (16) $N - 3 \leq N^* \leq N$.

The number of parameters is always reduced ($N^* < N$) after nondimensionalization as long as the matrix (16) contains nonzero entries, i.e. as long as not all the original parameters have been dimensionless already.

Example 1 (continued)**.** For the projectile problem, the matrix (16) is given by

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -2 & 0 & -1 \end{pmatrix} \tag{17}$$

which has rank 2. Hence the nullspace is 1-dimensional and there is only one dimensionless parameter.
The parameter ε given by (6) that we found in all scalings corresponds to the element $(-1, -1, 2)$ of the nullspace.
Of course the second row of the matrix does not contribute to the result, since obviously the fundamental dimension *mass* does not appear.

In this example, there are several ways to choose intrinsic reference units, which corresponds to the non-trivial null-space of the above matrix.

Note that for doing the "scaling", i.e. the mere nondimensionalization procedure, we actually do not need the complete model problem/equation, but only the list of variables and parameters:

Example 2. (The pendulum problem) We consider a point mass m at one end of a rigid, massless stick of length L. It freely rotates around a horizontal axis through the other end under the action of a fixed downward gravitational acceleration. Friction effects are neglected. We bring the pendulum into an initial, say horizontal, position, let it go, and are interested in the duration T^* of a complete period, returning to the starting position. The movement of the pendulum is described completely by the function $\theta^*(t^*)$, giving the angle between vertical lines and the pendulum. Parameters are the pendulum length L, the point mass m, and the gravitational acceleration g.

This leads to the following list of variables and parameters:

Quantity	Dimension	Unit
θ^*	–	1
t^*	Time	s
L	Length	m
m	Mass	kg
g	Acceleration	m s^{-2}

Table 2: variable - parameter list for example 2

The matrix (16) in this problem reads

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{18}$$

This matrix is invertible, which implies that there are no dimensionless parameters. The unique intrinsic reference time is given by $\sqrt{L/g}$. Thus the duration of a complete swing, i.e. one period, is given by

$$T^* = \sqrt{\frac{L}{g}}\, T\,, \tag{19}$$

where T is determined from the nondimensionalized problem, which however does not contain any parameters. Therefore T^* is a number which does not depend on the physical realization of the pendulum besides its length L. Instead of solving a mathematical problem, its value could also be determined from one experiment with known values of L and g. We also remark that the result is independent of the mass m.

1.2 Regular perturbations – asymptotic expansions

Example 1 (continued)**.** We recall the nondimensionalized projectile problem obtained by the Scaling (#3).

For small initial speed, i.e. $\varepsilon \ll 1$, it is a natural idea to look at the problem with $\varepsilon = 0$: For $\varepsilon = \mathbf{0}$ in (11)-(12), we get the **"reduced problem":**

$$x_0'' = -1 \tag{20}$$

$$x_0(0) = 0\,, \quad x_0'(0) = 1\,. \tag{21}$$

One can hope that the "reduced problem" (20)-(21) provides a good approximation for the "full problem", and that it is indeed "simpler" to solve.

For scaling (#3) this procedure seems to work since problem (20) is much easier to solve: integrating the constant twice and using the initial conditions (21) we easily obtain the **"reduced solution":**

$$x_0(t) = t - t^2/2. \tag{22}$$

However, this procedure of computing the "reduced solution" does *not* work for the other two scalings (#1) and (#2). For equation (4), setting $\varepsilon = 0$ completely changes the nature of the equation (from second order differential equation to algebraic equation) and the initial conditions lead to a contradiction; so the reduced problem does not even have a solution. For equation (8), when setting $\varepsilon = 0$, the initial position and velocity are zero, with a negative acceleration, so the reduced solution would have negative values for all positive times and thus not describe an upwards throw.
In these cases, the reduced problem for $\varepsilon = 0$ does not provide any approximation to the original problem, much less a good one. So for the purpose of asymptotics, the different scalings are not at all equivalent.

For analyzing approximations in the limit $\varepsilon \to 0$, one is working under the assumption that the dimensionless variables, and the derivatives appearing in the problem, are "of order 1" (an expression made more precise below, when we introduce the Landau symbols). The alternative non-dimensionalization choices (#1) and (#2) do not satisfy this assumption, in contrast to scaling (#3).
Calling the full problem an ε-perturbation of the "reduced problem", it seems obviously appropriate to call it a "singular perturbation" for scaling (#1), since for $\varepsilon = 0$ we lose the highest derivative.
For scaling (#3) it seems appropriate to call it a "regular perturbation", since the reduced problem is a simplification of the full problem that keeps the main features. In such a situation it makes sense to look for "small corrections" of the reduced problem for getting better approximations.

The above remark sketches the basic idea for "regular perturbations": take the "reduced solution" and add "correction terms" of increasing "order of ε" in an "asymptotic expansion". The resulting machinery of "perturbation theory" is widely and successfully used in physics, engineering,... We want to put it in a mathematically clear formulation with precise definitions, and in a few situations we can make the "formal asymptotics" rigorous, with proofs of convergence as $\varepsilon \to 0$.

In order to put the problem in an *abstract mathematical setting*, we define an operator F (which would contain the (Ordinary/Partial) Differential Equation(s) and all boundary/initial conditions in our context) :

$$F : B_1 \times [0, \varepsilon_0] \to B_2 \,, \tag{23}$$

where $(B_1, \|\cdot\|_1)$ and $(B_2, \|\cdot\|_2)$ are Banach spaces. With this definition, we formulate the abstract *perturbation problem*, i.e. a problem containing a small dimensionless parameter, as the equation

$$F(x_\varepsilon, \varepsilon) = 0 \,, \qquad 0 < \varepsilon \ll 1 \,. \tag{24}$$

We view (24) as the "exact" ("complete") problem, and define the *reduced problem* by setting $\varepsilon = 0$ in (24):

$$F(x_0, 0) = 0 \tag{25}$$

In case that (25) has a unique solution x_0, it is the *reduced solution*, and we hope that it is a meaningful approximation of the "exact" solution x_ε of (24).

Example 1 (continued)**.** In the example of the projectile problem, we pack the ODE for $x(t)$ and the 2 initial conditions into one function F. The problem consisting of (11)-(12) is hence reformulated by defining the function

$$F(x,\varepsilon) = \Big(x'' + (1+\varepsilon x)^{-2}, x(0), x'(0) - 1\Big) \tag{26}$$

mapping the real function x (belonging to some suitable function space) and the real number ε to a triple, whose first component is a real function, and the second and third components are real numbers.
A possible choice of spaces would be $B_1 = C^2([0,t_0])$, $B_2 = C([0,t_0])\times\mathbb{R}^2$ with some $t_0 > 0$.

Motivated by the example, we first consider the simple situation $F : \mathbb{R}^2 \to \mathbb{R}$ and the (algebraic) equation $F(x_\varepsilon,\varepsilon) = 0$ for the unknown number $x_\varepsilon \in \mathbb{R}$ and the small parameter $\varepsilon \in \mathbb{R}$. This situation is covered by the *implicit function theorem:* If

1. F is smooth enough,
2. the reduced problem $F(x_0,0) = 0$ has a solution $x_0 \in \mathbb{R}$, and
3. $\partial_x F(x_0,0) \neq 0$,

then there exists a smooth real function $\varepsilon \mapsto x_\varepsilon$, defined for $|\varepsilon|$ small enough, such that $F(x_\varepsilon,\varepsilon) = 0$ and x_0 is the value of this function at $\varepsilon = 0$.

Again, it seems reasonable to look for approximations x_0 for solutions x_ε by solving the *reduced problem* $F(x_0,0) = 0$. In the general case, this idea relies on a convergence and a continuity assumption, i.e. we assume

$$\lim_{\varepsilon\to 0} x_\varepsilon = x_0 \qquad \text{and} \qquad \lim_{\varepsilon\to 0} F(x_\varepsilon,\varepsilon) = F(x_0,0)\,, \tag{27}$$

where the first limit is in B_1 and the second in B_2.

Remaining on the formal side, we define:

Definition 1. *Let F be as in (23). Then $x_{as} \in B_1$, possibly depending on ε, is called a formal approximation for the solution of $F(x_\varepsilon,\varepsilon) = 0$, if the* ***residual*** *$r_\varepsilon := F(x_{as},\varepsilon)$ tends to zero in B_2 as $\varepsilon \to 0$.*

As a next step we aim to improve the approximation provided by the reduced solution x_0 by constructing a formal asymptotic expansion of x_ε in powers of ε at some finite order N, and write an "asymptotic expansion" as

$$x_\varepsilon = \sum_{n=0}^{N} \varepsilon^n x_n + O(\varepsilon^{N+1}) \qquad \text{as } \varepsilon \to 0\,, \tag{28}$$

where the maximal order N depends on the smoothness of F.
Here we use the *'big O'* symbol, one of the two **"Landau symbols"** (named after the famous physicist Lev Landau who heavily used "perturbation theory") , defined as

Definition 2. *Let $(B, \|\cdot\|)$ denote a Banach space, let $\varepsilon_0 > 0$, and let $y_\varepsilon \in B$ and $g_\varepsilon > 0$ for $0 < \varepsilon \leq \varepsilon_0$. Then we write*
1) ***"big O":*** *$y_\varepsilon = O(Cs)$ as $\varepsilon \to 0$, iff there exists a constant $C > 0$ such that $\|y_\varepsilon\| \leq C g_\varepsilon$ for $0 < \varepsilon \leq \varepsilon_0$.*
1b) $y_\varepsilon = O_s(g_\varepsilon)$ as $\varepsilon \to 0$, iff $y_\varepsilon = O(g_\varepsilon)$ and $g_\varepsilon = O(\|y_\varepsilon\|)$ as $\varepsilon \to 0$.
2) ***"little o":*** *$y_\varepsilon = o(g_\varepsilon)$ as $\varepsilon \to 0$, iff $\lim_{\varepsilon \to 0} \|y_\varepsilon\|/g_\varepsilon = 0$.*

Remark: in a more general definition, that we do not need, we take $g_\varepsilon \in B$, i.e. g_ε is not restricted real valued functions.

The following definition is a key to "perturbation theory".

Definition 3. *Let $(B, \|\cdot\|)$ denote a Banach space, let $\varepsilon_0 > 0$, $N \in \mathbb{N}$, let $x_0, \ldots, x_N \in B$ as well as $x_\varepsilon \in B$ for $0 < \varepsilon \leq \varepsilon_0$.*
The polynomial $x_0 + \varepsilon x_1 + \cdots + \varepsilon^N x_N$ is called an asymptotic expansion of order N for x_ε, iff

$$x_\varepsilon - \sum_{n=0}^{N} \varepsilon^n x_n = o(\varepsilon^N) \qquad \text{as } \varepsilon \to 0\,, \tag{29}$$

(which is slightly weaker than (28) where the remainder is $O(\varepsilon^{N+1})$).

The *formal formal aspect* is the *computation* of x_0 and, possibly, further coefficients $x_1, x_2, \ldots$ of the the expansion, which can require a lot of work and ideas. The *rigorous aspect* is the actual *proof* of the convergence in the limit, i.e. the proof of the*validity* of the expansion, which can require sophisticated mathematical tools.

In perturbation theory it is common to separate the two aspects, because they often require ideas and methods of different natures, and also because the rigorous aspect is often too difficult or impossible. Usually, the help of mathematician to find formal expansions is what the application requires. The situation is similar to numerical mathematics, where often the development of numerical methods is not accompanied by rigorous convergence results that are mostly impossible for the interesting nonlinear problems in real applications. Hence we shall concentrate on the formal aspect in the rest of this section.
Section 1.3

Example 1 (continued)**.** We return to the example of the projectile problem, and obtain the following:

With the approximation x_0 defined in (20), we obtain the residual

$$F(x_0, \varepsilon) = (-1 + (\varepsilon x_0 + 1)^{-2}, 0, 0) = \varepsilon \left(-\frac{x_0(2 + \varepsilon x_0)}{(\varepsilon x_0 + 1)^2}, 0, 0 \right)\,, \tag{30}$$

which converges to zero in $B_2 = C([0, t_0]) \times \mathbb{R}^2$ for every $t_0 > 0$.

As a next step we aim to improve the approximation by constructing a formal asymptotic expansion of x_ε for some suitable order $N > 0$, so we use the ansatz:

$$x_{as} = \sum_{n=0}^{N} \varepsilon^n x_n$$

For the computation of the order 1 asymptotic expansion $x_{as} = x_0 + \varepsilon x_1$ of a solution we assume sufficient regularity of F and compute formally

$$F(x_0 + \varepsilon x_1, \varepsilon) = F(x_0, 0) + \varepsilon \left(\partial_x F(x_0, 0) x_1 + \partial_\varepsilon F(x_0, 0) \right) + O(\varepsilon^2)\,, \tag{31}$$

where $\partial_x F$ and $\partial_\varepsilon F$ are partial Fréchet derivatives of F. Note that $\partial_\varepsilon F(x_0,0) \in B_2$, whereas $\partial_x F(x_0,0)$ is a linear map from B_1 to B_2. A $O(\varepsilon^2)$-residual can be produced by solving the linear problem

$$\partial_x F(x_0,0)x_1 = -\partial_\varepsilon F(x_0,0)$$

for x_1. Unique solvability is guaranteed, if there exists the inverse $\partial_x F(x_0,0)^{-1} : B_2 \to B_1$. For higher order expansions the general computations are somewhat involved and we only state the general result:

Lemma 4. *Let, for fixed $x \in B_1$, F possess a formal asymptotic expansion*

$$F(x,\varepsilon) = \sum_{k=0}^{N} F_k(x)\varepsilon^k + O\left(\varepsilon^{N+1}\right) \tag{32}$$

with $(N+1)$-times Fréchet differentiable coefficients F_k, and let $x_\varepsilon \in B_1$ have an asymptotic expansion

$$x_\varepsilon = \sum_{k=0}^{N} x_k\varepsilon^k + O\left(\varepsilon^{N+1}\right) . \tag{33}$$

Then $F(x_\varepsilon,\varepsilon)$ has a formal asymptotic expansion of the form

$$F(x_\varepsilon,\varepsilon) = F(x_0,0) + \sum_{k=1}^{N} \left(\partial_x F_0(x_0)x_k + G_k(x_0,\ldots,x_{k-1})\right)\varepsilon^k + O(\varepsilon^{N+1}) . \tag{34}$$

With an nth order asymptotic expansion of the solution, a $O(\varepsilon^{N+1})$-residual can be produced by consecutively solving the problems

$$\partial_x F_0(x_0)x_k = -G_k(x_0,\ldots,x_{k-1}) , \qquad k = 1,\ldots,N , \tag{35}$$

which are all linearized versions of the leading order problem $F(x_0,0) = 0$.

Definition 5. *Let $F : (B_1, \|\cdot\|_1) \times [0,\varepsilon_0] \to (B_2, \|\cdot\|_2)$ be continuously Fréchet differentiable. Then the problem $F(x_\varepsilon,\varepsilon) = 0$ is called regularly perturbed, if the reduced problem $F(x_0,0) = 0$ has a solution $x_0 \in B_1$, such that the Fréchet derivative $\partial_x F(x_0,0) : B_1 \to B_2$ has a bounded inverse. Otherwise the problem is called singularly perturbed.*

The bounded invertibility of the Fréchet derivative, which is equal to the linearization of F, is the generalization of the above mentioned condition $\partial_x F(x_0,0) \neq 0$ for the application of the implicit function theorem in the case $B_1 = B_2 = \mathbb{R}$.

Example 1 (continued)**.** We return to the scaled projectile problem (11)- (12):

$$\begin{aligned} x'' &= -(\varepsilon x + 1)^{-2} , \\ x(0) &= 0 , \qquad x'(0) = 1 . \end{aligned} \tag{36}$$

Substituting the expansion - i.e. "plugging the ansatz

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + O\left(\varepsilon^3\right) \tag{37}$$

into the equation" yields

$$\begin{aligned} x_0'' + \varepsilon x_1'' + \varepsilon^2 x_2'' + O\left(\varepsilon^3\right) &= -\left(1 + \varepsilon x_0 + \varepsilon^2 x_1 + \varepsilon^3 x_2 + O\left(\varepsilon^4\right)\right)^{-2} , \\ x_0(0) + \varepsilon x_1(0) + \varepsilon^2 x_2(0) + O\left(\varepsilon^3\right) &= 0 , \\ x_0'(0) + \varepsilon x_1'(0) + \varepsilon^2 x_2'(0) + O\left(\varepsilon^3\right) &= 1 . \end{aligned} \tag{38}$$

We use the Taylor expansion

$$-\left(1+\varepsilon x_0+\varepsilon^2 x_1+\varepsilon^3 x_2+O(\varepsilon^4)\right)^{-2} = -1+\varepsilon 2x_0+\varepsilon^2(2x_1-3x_0^2)+O(\varepsilon^3) \tag{39}$$

of the right hand side of the differential equation and compare coefficients of ε^0:

$$x_0'' = -1\,, \qquad x_0(0)=0\,, \quad x_0'(0)=1\,. \tag{40}$$

Coefficients of ε^1:

$$x_1'' = 2x_0\,, \qquad x_1(0)=x_1'(0)=0\,. \tag{41}$$

Coefficients of ε^2:

$$x_2'' = 2x_1-3x_0^2\,, \qquad x_2(0)=x_2'(0)=0\,. \tag{42}$$

The reduced problem has been solved already:

$$x_0(t) = t-\frac{t^2}{2}\,. \tag{43}$$

For the higher order coefficients we obtain

$$x_1(t) = \frac{t^3}{3}\left(1-\frac{1}{4}t\right)\,, \qquad x_2(t) = -\frac{t^4}{4}\left(1-\frac{11}{15}t+\frac{11}{90}t^2\right)\,, \tag{44}$$

and, thus, the formal expansion

$$x(t) = t-\frac{t^2}{2}+\varepsilon\frac{t^3}{3}\left(1-\frac{1}{4}t\right)-\varepsilon^2\frac{t^4}{4}\left(1-\frac{11}{15}t+\frac{11}{90}t^2\right)+O(\varepsilon^3)\,. \tag{45}$$

For finding an approximation of the time T, where the projectile hits the ground, we use the same procedure and make the ansatz

$$T = T_0+\varepsilon T_1+O(\varepsilon^2)\,. \tag{46}$$

Substitution into the expansion for x and re-expansion gives

$$x(T) = \frac{T_0}{2}(2-T_0)+\varepsilon\left((1-T_0)T_1+\frac{T_0^3}{3}\left(1-\frac{T_0}{4}\right)\right)+O(\varepsilon^2)\,, \tag{47}$$

and comparing coefficients of powers of ε in the equation $x(T)=0$ leads to $T_0=2$, $T_1=4/3$, and therefore

$$T = 2+\frac{4}{3}\varepsilon+O(\varepsilon^2)\,. \tag{48}$$

From physical intuition, a positive correction to the leading order term is what we expect, since the full problem yields a weaker gravitation than the leading order approximation with the constant gravitation given at the surface of the earth, such that the projectile can be expected to fly higher and longer.

Asymptotic Expansion $\varepsilon \to 0$ with N fix vs. **Taylor Expansion**: $N \to \infty$ with ε fix

At first sight these 2 concepts might seem identical (and are sometimes mixed up by physicists), especially since in practice we often use Taylor expansions to construct asymptotic expansions. The fundamental difference is that with asymptotic series we are interested in the expansion in the limit $\varepsilon \to 0$ at fixed order N, whereas with Taylor series we are interested in the convergence of the infinite series as $N \to \infty$ for fixed ε.
There are functions for which asymptotic expansions up to arbitrary order exist, although the corresponding infinite series diverge for all $\varepsilon > 0$.

We give another definition in this context:

Definition 6. ***Asymptotic equivalence***
For $0 < \varepsilon \leq \varepsilon_0$ let x_ε and y_ε be ε-dependent elements of a normed space $(B, \|\cdot\|)$. Then we call x_ε and y_ε ***asymptotically equivalent for*** *$\varepsilon \to 0$, if*

$$\|x_\varepsilon - y_\varepsilon\| = o(\|y_\varepsilon\|) \qquad \text{for} \quad \varepsilon \to 0 \tag{49}$$

i.e. if the relative error of the approximation of x_ε by y_ε (or the other way round) goes to zero when $\varepsilon \to 0$.

The name „equivalence“ is justified since obviously the above definition is symmetric w.r.t. x_ε and y_ε. Taking y_ε as an approximation for x_ε, it is common to call y_ε an *asymptotic approximation* for x_ε.

Example for asymptotic approximation/equivalence:
Let $B = \mathbb{R}$ and $x_\varepsilon = x(\varepsilon) = \sin(\varepsilon)$. Then the following 3 functions

$$y(\varepsilon) = \varepsilon, \qquad y(\varepsilon) = \varepsilon + 2\varepsilon^2, \qquad y(\varepsilon) = \varepsilon - \frac{1}{6}\varepsilon^3 \tag{50}$$

are each an asymptotic approximation of $x(\varepsilon)$, since we have $x(\varepsilon) = \varepsilon - \frac{1}{6}\varepsilon^3 + O(\varepsilon^5)$.

Example 3. "function with boundary layer at x=0"
Let $B = C([0,1])$ and consider the real function $x_\varepsilon \in B$, $x_\varepsilon : [0,1] \to \mathbb{R}$ given by

$$x_\varepsilon(t) = t + \exp\left(-\frac{t}{\varepsilon}\right), \qquad 0 \leq t \leq 1\,.$$

For positive $t > 0$ the sequence of functions $x_\varepsilon(t)$ converges with $\varepsilon \to 0$ towards $x_0(t) = t$. For $t = 0$, however, the value 1 is an asymptotic approximation - see fig 1.

a) If we consider x_ε as an element of the Banach-space $C([0,1])$ with norm

$$\|x\|_\infty = \max_{t\in[0,1]} |x(t)|\,, \tag{51}$$

then x_0 is not asymptotically equivalent to x_ε, since we have

$$\|x_\varepsilon - x_0\|_\infty = x_\varepsilon(0) - x_0(0) = 1\,, \qquad \|x_0\|_\infty = 1 \tag{52}$$

Hence there is no ε independent asymptotic approximation for x_ε.

b) If, however, we choose $B = L^p((0,1))$, with $1 \leq p < \infty$, as the underlying Banach-space for this function $x_\varepsilon(t)$, the situation is completely different ! Now we use an integral-norm

$$\|f\|_p = \left(\int_0^1 |f(t)|^p dt\right)^{1/p}. \tag{53}$$

And for $p = 1$ we have

$$\|x_\varepsilon - x_0\|_1 = \int_0^1 e^{-\frac{t}{\varepsilon}} dt = -\varepsilon e^{-\frac{t}{\varepsilon}}\Big|_0^1 \leq \varepsilon\,, \qquad \|x_0\|_1 = \frac{1}{2}\,.$$

Now x_0 is an asymptotic approximation of x_ε, in the sense of the L^1-Norm.
The point is that the quality of an approximation is "poorer" for integral norms than for pointwise norms. □

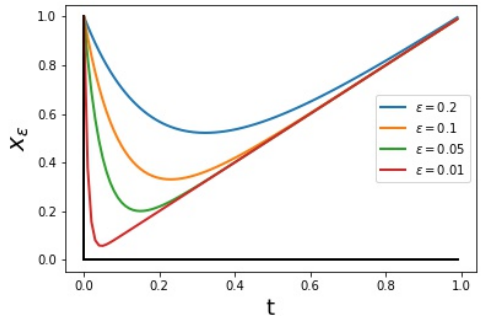

Figure 1: The function $x_\varepsilon(t)$ in $[0,1]$ for $\varepsilon = 0.2, 0.1, 0.05, 0.01$

1.3 Validity of formal asymptotic approximations

We now turn to the task of proving rigorously that the construction of a formal asymptotic expansion actually yields a valid approximation to the solution of the original problem *if the problem is regularly perturbed.*

As already mentioned, this situation is covered by an infinite-dimensional version of the implicit function theorem. In preparation, we need some basic results from calculus on Banach spaces (denoted by B_1, B_2 etc.) that we state without proof:

Theorem 7 (Mean value theorem)**.** *Suppose $U \subset B_1$ is open and convex and $F : U \to B_2$ is C^1 (i.e. continuously Fréchet differentiable). Then*

$$\|F(y) - F(x)\| \leq \sup_{t\in[0,1]} \|DF[(1-t)x + ty]\| \|y - x\| \quad \forall x, y \in U. \tag{54}$$

□

Denoting the set of continuous linear maps from B_1 to B_2 by $L(B_1, B_2)$, we also have

Lemma 8. *The set $GL(B_1, B_2)$ of linear isomorphisms from B_1 to B_2 is open in $L(B_1, B_2)$.*

□

Recall also one of the pillars of functional analysis,

Theorem 9 (Open mapping theorem)**.** *Let the linear mapping $A \in L(B_1, B_2)$ be onto. Then A is an open mapping (i.e. A maps open sets into open sets).*

□

As it is frequently done in (courses on) multivariable calculus, we first prove a result on *inverse mappings,* utilizing Banach's fixed point theorem:

Theorem 10 (Inverse mapping theorem)**.** *Let $F : U \subset B_1 \to B_2$ be C^1, $x_0 \in U$ and suppose $DF(x_0)$ is a linear isomorphism. Then F is a C^1 diffeomorphism (i.e. bijective with C^1 inverse) of some neighbourhood of x_0 onto some neighbourhood of $F(x_0)$.*

Proof. We begin with the observation that the affine transformation

$$F(x_0 + x) \mapsto DF(x_0)^{-1} [F(x_0 + x) - F(x_0)]$$

is invertible (and smooth). Therefore, it suffices to prove the theorem under the simplifying assumptions $B_1 = B_2$, $x_0 = 0$, $F(x_0) = 0$ and $DF(x_0) = I$ (identity).

In the next step, we want to prove the *existence* of an inverse. To do this, we reformulate the equation

$$F(x) = y \tag{55}$$

(which should have a unique solution in a neighbourhood of x_0) as a fixed point problem. Rewriting (55) as $y + x - F(x) = x$, we are therefore motivated to introduce the family of mappings

$$G_y(x) := y + x - F(x). \tag{56}$$

If G_y is a contraction in some neighbourhood of $x_0 = 0$ for suitable y, the Banach fixed point theorem will immediately yield the existence of the desired inverse. Thus, checking the properties of G_y, we get

1. $DG_y(0) = I - DF(0) = 0$, so by continuity there is a $\delta > 0$ such that

$$\|DG_y(x)\| = \|I - DF(x)\| \leq 1/2 \tag{57}$$

   for all $x \in K_\delta(0) := \{x \in B_1 | \|x\| \leq \delta\}$.

2. By (54),

$$\begin{aligned} \|x - F(x)\| &= \|x - F(x) - [0 - F(0)]\| \\ &\leq \sup_{t\in[0,1]} \|I - DF(tx)\| \|x - 0\| \leq \frac{1}{2}\|x\| \leq \frac{\delta}{2} \end{aligned} \tag{58}$$

   on $K_\delta(0)$.

3. For any $y \in K_{\delta/2}(0)$, $x \in K_\delta(0)$,

$$\|G_y(x)\| \leq \|y\| + \|x - F(x)\| \leq \delta, \tag{59}$$

   which tells us that G_y for such y maps the closed ball $K_\delta(0)$ into itself.

4. Again by (54), we have

$$\|G_y(x_1) - G_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|, \tag{60}$$

   and we can conclude that G_y is a *contraction on* $K_\delta(0)$ under the stated conditions.

Thus, for every $y \in K_{\delta/2}(0)$, G_y has a unique fixed point in $K_\delta(0)$. Since $G_y(x) = x \iff F(x) = y$, we have therefore proven the existence of an inverse $F^{-1} : K_{\delta/2}(0) \to F^{-1}(K_{\delta/2}(0)) \subset K_\delta(0)$. It remains to investigate the continuity and smoothness of F^{-1}:

By (60) with $y = 0$,

$$\|x_1 - F(x_1) - [x_2 - F(x_2)]\| \leq \frac{1}{2}\|x_1 - x_2\|, \tag{61}$$

so

$$|\|x_1 - x_2\| - \|F(x_1) - F(x_2)\|| \leq \frac{1}{2}\|x_1 - x_2\|, \tag{62}$$

implying

$$\|x_1 - x_2\| \leq 2\|F(x_1) - F(x_2)\| \tag{63}$$

and consequently

$$\|F^{-1}(y_1) - F^{-1}(y_2)\| \leq 2\|y_1 - y_2\|. \tag{64}$$

This proves continuity of F^{-1}.

Regarding differentiability, recall that the Fréchet derivative is defined in exactly the same way as the (total) derivative in multivariable calculus. Now, by Lemma 8, we can choose $\delta > 0$ such that $DF(x)^{-1}$ exists on $K_\delta(0)$. As a consequence of Theorem 9, $DF(x)^{-1}$ is continuous and $\|DF(x)^{-1}\| \leq M$ for some M (for all $x \in K_\delta(0)$). For $y_1, y_2 \in K_{\delta/2}(0)$, $x_1 = F^{-1}(y_1)$, $x_2 = F^{-1}(y_2)$, it therefore holds

$$\begin{aligned}
&\|F^{-1}(y_1) - F^{-1}(y_2) - DF(x_2)^{-1}(y_1 - y_2)\| \\
&= \|x_1 - x_2 - DF(x_2)^{-1}\left[F(x_1) - F(x_2)\right]\| \\
&= \|DF(x_2)^{-1}\left[DF(x_2)(x_1 - x_2) - (F(x_1) - F(x_2))\right]\| \\
&\leq M\|F(x_1) - F(x_2) - DF(x_2)(x_1 - x_2)\|.
\end{aligned} \tag{65}$$

This shows that F^{-1} is indeed Fréchet differentiable, with a continuous derivative given by $DF(F^{-1}(y))^{-1}$. □

With the inverse mapping theorem, it is straightforward to prove

Theorem 11 (Implicit function theorem)**.** *Let $U \subset B_1$, $V \subset B_2$ be open and $F : U \times V \to B_3$ be C^1. At some point $(x_0, y_0) \in U \times V$, assume that the derivative of F with respect to y, $\partial_y F(x_0, y_0) : B_2 \to B_3$, is invertible. Then there exist neighbourhoods U_0 of x_0 and W_0 of $F(x_0, y_0)$ and a unique C^1 map $H : U_0 \times W_0 \to V$ such that, for all $(x, w) \in U_0 \times W_0$,*

$$F(x, H(x, w)) = w. \tag{66}$$

Proof. Consider $\Phi : U \times V \to B_1 \times B_3$, $\Phi(x, y) = (x, F(x, y))$. The derivative $D\Phi(x_0, y_0)$ reads

$$D\Phi(x_0, y_0)(x_1, y_1) = \begin{pmatrix} I & 0 \\ \partial_x F(x_0, y_0) & \partial_y F(x_0, y_0) \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \tag{67}$$

which is invertible since $\partial_y F(x_0, y_0)$ is (what does the inverse look like?). Thus, Φ has a local C^1 inverse

$$\Phi^{-1} : U_0 \times W_0 \to U \times V$$

and, writing

$$\Phi^{-1}(x, w) = (x, H(x, w)), \tag{68}$$

the H such defined has the desired properties. □

Before we employ the implicit function theorem to prove convergence of regular perturbations to their asymptotic approximations, we need to address a technicality:

We defined our abstract perturbation problem $F(x_\varepsilon, \varepsilon) = 0$ on the *closed* ε-interval $[0, \varepsilon_0]$. Therefore, the derivative $\partial_\varepsilon F(x, \varepsilon)$ at $\varepsilon = 0$ (whose existence, not invertibility

we need) is defined only in a one-sided sense and it is not clear without further analysis whether the results just obtained fully apply. However, in a regular perturbation problem derived from a scaling procedure, we can safely assume that the properties listed in Definition 5 will also hold when we extend our domain of definition to a set of the form $B_1 \times (-\varepsilon', \varepsilon'')$, with ε' arbitrarily small. Keeping this in mind, we can prove our main result:

Theorem 12. *If the problem $F(x_\varepsilon, \varepsilon) = 0$, $F : B_1 \times (-\varepsilon', \varepsilon'') \to B_2$ is regularly perturbed, with reduced solution $x_0 \in B_1$, there is an $\varepsilon_1 > 0$ such that*

1. *there is a unique solution $x_\varepsilon \in B_1$ for all $\varepsilon \in [0, \varepsilon_1]$ and*
2. $\lim_{\varepsilon \to 0} \|x_\varepsilon - x_0\| = 0$.

The same result holds for any asymptotic approximation.

Proof. Apply the implicit function theorem with the replacements $B_1 \mapsto (-\varepsilon', \varepsilon'')$, $B_2 \mapsto B_1$ and $B_3 \mapsto B_2$ and consider $H(\varepsilon, 0)$. □

Remark. Theorem 12 provides a fairly general justification for the use of asymptotic expansions in the context of a regular, "well-behaved" perturbation. In scientific practice, however, most problems of interest are *not* regular and different, far more advanced techniques are required. Furthermore, even when studying a problem as simple as Example 1, it might be more insightful to, e.g., directly calculate estimates on $\|x_\varepsilon - x_0\|$ instead of proving that the requirements for Theorem 12 to apply are met.

1.4 Singular perturbations - boundary layers

Even in case that we perturb the highest derivative in the problem/equation and hence expect a singular perturbation, we can still use the idea of an asymptotic expansion with the reduced solution as the basis term. We can (sometimes) save the idea by adding extra "boundary layer terms" that, e.g., remedy the problem that the reduced solution can not fulfill all boundary/initial conditions.

Example 4. *Consider the initial value problem*

$$\varepsilon u' = -u + x + \varepsilon\,, \qquad u(0) = 1\,. \tag{69}$$

Since the reduced problem

$$0 = -u_0 + x\,, \qquad u_0(0) = 1\,, \tag{70}$$

has no solution, the problem is singularly perturbed.

However, our previously studied function

$$u_\varepsilon(x) = e^{-x/\varepsilon} + x \tag{71}$$

is the solution of (69). And we see from Fig 1. that the solution $\overline{u}(x) = x$ of the reduced differential equation (70) approximates the exact solution everywhere except close to $x = 0$.

In order to develop this idea into a well defined mathematical machinery where we introduce "(regularizing) local / layer variables" at (sets of) points like the point $x = 0$ in the above example, and then systematically find "boundary layer terms" that fix the problem at these (sets of) points, we need some clever but tedious definitions.

Definition 13. *Let $D \subset \mathbb{R}^n$ be a domain and let $u_\varepsilon \in C(\overline{D})$ for $0 < \varepsilon \leq \varepsilon_0$.*
a) The family/sequence of functions $\{u_\varepsilon : 0 < \varepsilon \leq \varepsilon_0\}$ is called regular in D, if u_ε converges uniformly, i.e. with respect to the supremum norm $\|u\|_D := \sup_{x\in D} |u(x)|$, in D as $\varepsilon \to 0$.
b) Let $S \subset \overline{D}$ be a C^1-manifold with dimension smaller than n. The family $\{u_\varepsilon\}$ has layer behavior at S if it is not regular in D, but regular in each D_1 with $\overline{D}_1 \subset \overline{D} \setminus S$.

Lemma 14. *Let $\{u_\varepsilon\}$ have layer behavior at S, then there exists an outer limit $\overline{u} \in C(\overline{D} \setminus S)$ such that*

$$\lim_{\varepsilon\to 0} \|u_\varepsilon - \overline{u}\|_{D_1} = 0\,, \qquad \text{for } \overline{D}_1 \subset \overline{D} \setminus S\,. \tag{72}$$

Proof: Since $\overline{D} \setminus S$ is open relative to $\overline{D}$, every point in $\overline{D} \setminus S$ lies in a suitable D_1. This implies the existence of a pointwise limit $\overline{u}$, which is continuous as a consequence of the uniform convergence in D_1. $\square$

Example 4 (continued)**.** In example 4,

$$u_\varepsilon(x) = e^{-x/\varepsilon} + x\,, \qquad D = (0,1)\,, \quad S = \{0\}\,, \quad \overline{u}(x) = x\,. \tag{73}$$

Example 5.

$$u_\varepsilon(x,y) = \tanh\frac{x-y^2}{\varepsilon}, \qquad D = \mathbb{R}^2\,, \quad S = \{(x,y)\,\big|\, x - y^2 = 0\}\,, \tag{74}$$

$$\overline{u}(x,y) = \begin{cases} -1 & \text{für} \quad x - y^2 < 0\,, \\ 1 & \text{für} \quad x - y^2 > 0\,. \end{cases} \tag{75}$$

The function $\overline{u}$ describes the behavior of u_ε "away" from layers.
For studying u_ε inside a boundary layer, i.e. close to S, we use a mathematical "magnifying glass" in the form of a coordinate transformation: For a C^1-manifold S with $\dim(S) = n-k$ we introduce "local coordinates" $z = z(x)$, defined in a neighborhood of S, such that

$$S = \{x : z_1(x) = \cdots = z_k(x) = 0\}\,. \tag{76}$$

Local variables $\xi(x)$ close to S are then given by a "rescaling with powers of ε":

$$\begin{aligned} \xi_i &= z_i \varepsilon^{-\alpha_i}\,, \qquad \text{with } \alpha_i > 0\,, \text{ for } i = 1,\ldots,k\,, \\ \xi_i &= z_i\,, \qquad \text{for } i = k+1,\ldots,n\,. \end{aligned}$$

Local variables serve their purpose if they are *regularizing*, i.e. if the transformed family

$$U_\varepsilon(\xi) := u_\varepsilon\left(x(\xi,\varepsilon)\right) \tag{77}$$

is regular in an appropriate region.

Example 4 (continued)**.** For $u_\varepsilon(x) = e^{-x/\varepsilon} + x$, a local variable close to $x = 0$ is given by $\xi = x\varepsilon^{-\alpha}$. The transformed function is

$$U_\varepsilon(\xi) = \exp(-\xi\varepsilon^{\alpha-1}) + \varepsilon^\alpha \xi$$

For $\alpha \geq 1$, it is regular on ξ-intervals of the form $(0, X)$ with

$$\lim_{\varepsilon\to 0} U_\varepsilon(\xi) = \begin{cases} e^{-\xi} & \text{für} \quad \alpha = 1\,, \\ 1 & \text{für} \quad \alpha > 1\,, \end{cases}$$

and for $0 < \alpha < 1$, it is regular on intervals of the form (X_1, X_2) with $X_1 > 0$, and $\lim_{\varepsilon\to 0} U_\varepsilon(\xi) = 0$.

Example 6. Not always a reasonable regularization can be achieved by local variables. In the previous section we have been interested in the behavior of functions like

$$u_\varepsilon(t) = e^{-\varepsilon t} \sin t$$

as $t \to \infty$. This is equivalent to studying the local behavior of $s = 1/t$ near $s = 0$ of

$$v_\varepsilon(s) = e^{-\varepsilon/s} \sin \frac{1}{s} \,.$$

Obviously $\{v_\varepsilon\}$ has layer behavior with the outer limit $\overline{v}(s) = \sin(1/s)$. The local variable $\sigma = s\varepsilon^{-\alpha}$ leads to

$$V_\varepsilon(\sigma) = \exp(-\varepsilon^{1-\alpha}/\sigma) \sin\left(\frac{1}{\varepsilon^\alpha \sigma}\right) .$$

For $0 < \alpha \leq 1$ there is no σ-interval, where V_ε is regular. For $\alpha > 1$ we have $\lim_{\varepsilon\to 0} V_\varepsilon(\sigma) = 0 = v_\varepsilon(0)$, which is not very informative.

In the example 4, $u_\varepsilon(x) = e^{-x/\varepsilon} + x$, the local variable $\xi = x/\varepsilon$ seems most appropriate, since the corresponding local approximation $e^{-\xi}$ is the "most informative". The other local approximations, 0 and 1, can be recovered from it with $\xi = 0$ and $\xi \to \infty$. These ideas are formalized in the following.

Definition 15.
a) Let ξ be a local variable. Then we define the local limit of u_ε with respect to ξ by

$$\left(\lim_{\xi} u_\varepsilon\right)(\xi) := \lim_{\varepsilon\to 0} u_\varepsilon(x(\xi,\varepsilon)) \,. \tag{78}$$

The domain of $\lim_\xi u_\varepsilon$ is the union of all compact sets where the above limit is uniform.

b) Let ξ_1 and ξ_2 be local variables, and let D_2 and D_{12} be the domains of the local limits of u_ε and, respectively, of $\lim_{\xi_1} u_\varepsilon$ with respect to ξ_2. Then the local limit of u_ε with respect to ξ_2 is contained in the local limit with respect to ξ_1, if $D_2 \subset D_{12}$ and

$$\lim_{\xi_2}\left(\lim_{\xi_1} u_\varepsilon\right) = \lim_{\xi_2} u_\varepsilon \,, \tag{79}$$

uniformly in compact subsets of D_2.

c) A local limit is called significant, if it is not contained in any other local limit, i.e. if it is a maximal element with respect to the order relation defined in b).
Local variables corresponding to significant limits are called layer variables.

Example 7. There can be more than one significant limit. Consider

$$u_\varepsilon(x) = x\left(\frac{e^{-x/\varepsilon}}{x+\varepsilon^2} + 1\right) \qquad 0 \leq x \leq 1 \,. \tag{80}$$

Obviously $\overline{u}(x) = x$, and $\{u_\varepsilon\}$ has layer behavior since $\lim_{\varepsilon\to 0} u_\varepsilon(\varepsilon^2) = 1/2$. It is easily seen that there are two layer variables: $\xi_1 = x/\varepsilon$ and $\xi_2 = x/\varepsilon^2$. The corresponding local limits are $e^{-\xi_1}$ and, respectively, $\frac{\xi_2}{1+\xi_2}$.

We propose a procedure for finding approximations which are uniformly valid within and away from layer regions. We consider the situation with exactly one layer variable ξ. The function

$$\widetilde{u}_\varepsilon := u_\varepsilon - \overline{u}$$

is small away from S. We shall use the *heuristic principle* that $\widetilde{u}_\varepsilon$ can be approximated uniformly by the significant local limit. This means that

$$\overline{u} + \lim_{\xi}(u_\varepsilon - \overline{u})$$

is expected to be a uniform asymptotic approximation for u_ε. The second term, which is small away from S, is called a *layer correction.*

Example 7 (continued)**.** The above procedure can be iterated in the case of more than one layer variables. We return to example (80) and approximate u_ε by

$$\overline{u} + \lim_{\xi_1}(u_\varepsilon - \overline{u}) + \lim_{\xi_2}\left(u_\varepsilon - \overline{u} - \lim_{\xi_1}(u_\varepsilon - \overline{u})\right) .$$

The three contributions depend on the variables x, ξ_1, and, respectively, ξ_2. Straightforward computations show

$$\left\| u_\varepsilon - x - e^{-x/\varepsilon} + \frac{1}{1 + x/\varepsilon^2} \right\|_{(0,1)} = O(\varepsilon) . \tag{81}$$

The approximation can be improved by dividing the error by ε and repeating the procedure. Iteration leads to an asymptotic expansion for u_ε in the form

$$u_\varepsilon(x) = \sum_{k=0}^{N} \left(\overline{u}_k(x) + \widetilde{u}_k\left(\frac{x}{\varepsilon}\right) + \widehat{u}_k\left(\frac{x}{\varepsilon^2}\right)\right) \varepsilon^k + O(\varepsilon^{N+1}) . \tag{82}$$

The coefficients of ε^0 have already appeared above:

$$\overline{u}_0(x) = \overline{u}(x) = x\,, \qquad \widetilde{u}_0(\xi_1) = e^{-\xi_1}\,, \qquad \widehat{u}_0(\xi_2) = -\frac{1}{1+\xi_2}\,. \tag{83}$$

For the coefficients of ε^1 we obtain

$$\overline{u}_1(x) = 0\,, \qquad \widetilde{u}_1(\xi_1) = \frac{1 - e^{-\xi_1}}{\xi_1}\,, \qquad \widehat{u}_1(\xi_2) = -\frac{1}{1+\xi_2}\,. \tag{84}$$

The fact that the layer corrections do not contribute outside of the layer can be written as

$$\lim_{\xi_1\to\infty} \widetilde{u}_k(\xi_1) = \lim_{\xi_2\to\infty} \widehat{u}_k(\xi_2) = 0\,. \tag{85}$$

We now turn our attention to problems whose solutions have layer behavior.

Example 8. Consider the boundary value problem

$$\begin{aligned} -\varepsilon u'' + u' + u = 0\,, \qquad & 0 < x < 1\,, \\ u(0) = 1\,, \qquad & u(1) = 0\,. \end{aligned} \tag{86}$$

This problem is singularly perturbed since solutions of the reduced differential equation

$$u' + u = 0\,, \tag{87}$$

cannot satisfy both boundary conditions.

We shall call a differential equation containing a small parameter ε *singularly perturbed,* if setting $\varepsilon = 0$ changes the *type* of the equation. By 'type' we mean the class of auxiliary conditions which, together with the differential equation, produce a well posed problem. The type of an ODE is determined by its order. Thus, an ODE is singularly perturbed, if the reduced equation is of lower order. For PDEs a definition of type is a more subtle matter.

Problems consisting of singularly perturbed differential equations and auxiliary (i.e. initial or boundary) conditions are typically singularly perturbed problems. Solutions of the reduced differential equation typically

1. cannot satisfy all auxiliary conditions and/or
2. do not have the smoothness properties expected from the exact solution of the perturbed problem.

Heuristic rule: *Solutions of problems involving singularly perturbed differential equations have layer behavior at manifolds, where auxiliary conditions are posed or where a loss of smoothness occurs.*

As a consequence, the outer limit of the solution can be chosen as a solution of the reduced differential equation.

Example 8 (continued)**.** The general solution of the reduced ODE (87) is given by

$$\overline{u}(x) = Ae^{-x}\,. \tag{88}$$

It is not clear yet how the constant A has to be chosen. According to the heuristic assumption boundary layers may occur at $x = 0$ and/or $x = 1$. In terms of the local variable $\xi = x\varepsilon^{-\alpha}$ near $x = 0$, the differential equation takes the form

$$-\varepsilon^{1-2\alpha}\ddot{u} + \varepsilon^{-\alpha}\dot{u} + u = 0\,, \qquad \left(\dot{u} = \frac{du}{d\xi}\right)\,. \tag{89}$$

For passing to the limit $\varepsilon \to 0$ we need to multiply by a power of ε in dependence of α. In the limit we obtain

$$\begin{aligned} \dot{u} &= 0 \qquad \text{for } 0 < \alpha < 1\,,\\ -\ddot{u} + \dot{u} &= 0 \qquad \text{for } \alpha = 1\,,\\ -\ddot{u} &= 0 \qquad \text{for } \alpha > 1\,. \end{aligned}$$

Definition 16.
a) Let L_ε be a linear differential operator, which becomes $\mathcal{L}_\varepsilon$ when written in terms of the local variable ξ. Let $\gamma \in \mathbb{R}$ be chosen such that $\varepsilon^\gamma \mathcal{L}_\varepsilon(u) = O_s(1)$. Then the local degeneration of L_ε with respect to ξ is defined by

$$\lim_{\xi} L_\varepsilon(u) := \lim_{\varepsilon\to 0} \varepsilon^\gamma \mathcal{L}_\varepsilon(u)\,.$$

b) The local degeneration of L_ε with respect to ξ_2 is contained in the local degeneration with respect to ξ_1, iff

$$\lim_{\xi_2}\left(\lim_{\xi_1} L_\varepsilon(u)\right) = \lim_{\xi_2} L_\varepsilon(u)\,.$$

A significant degeneration is a local degeneration, which is not contained in any other.

Correspondence principle: *When u_ε solves a linear differential equation $L_\varepsilon(u_\varepsilon) = 0$, then we expect that every significant limit of u_ε solves the corresponding significant degeneration of the differential equation.*

Example 8 (continued)**.** In the example (89) the local variable $\xi = x/\varepsilon$ leads to a significant degeneration near $x = 0$. For a boundary layer near $x = 1$ we find similarly $\eta = (1-x)/\varepsilon$. The above heuristic principles lead to the ansatz

$$u(x,\varepsilon) = \overline{u}(x) + \widetilde{u}(\xi) + \widehat{u}(\eta) + O(\varepsilon)\,, \tag{90}$$

with

$$\lim_{\xi\to\infty} \widetilde{u}(\xi) = \lim_{\eta\to\infty} \widehat{u}(\eta) = 0\,. \tag{91}$$

Actually we shall need the stronger assumption that also the derivatives of $\widetilde{u}$ and of $\widehat{u}$ tend to zero as the arguments tend to infinity. We obtain the equations for $\overline{u}$, $\widetilde{u}$ and $\widehat{u}$ by inserting the ansatz into (86) and then writing the resulting expression solely in terms of x, ξ and η. Passing to the limit $\varepsilon \to 0$, we get for $\overline{u}$:

$$\overline{u}_x + \overline{u} = 0$$

so we get back the reduced equation (87). Now we write the whole equation in terms of the independent variable ξ, i.e we set $x = \varepsilon\xi$ and $\eta = \frac{1}{\varepsilon} - \xi$. Passing to the limit $\varepsilon \to 0$ and using the decay assumption on $\widehat{u}$ implies

$$-\frac{d\widetilde{u}^2}{d\xi^2} + \frac{d\widetilde{u}}{d\xi} = 0\,,$$

which is the significant degeneration of the problem corresponding to the boundary layer $\{x = 0\}$. Therefore $\widetilde{u} = 0$, since this is the only decaying solution. The analogous procedure with the independent variable η leads to

$$-\frac{d\widehat{u}^2}{d\eta^2} - \frac{d\widehat{u}}{d\eta} = 0\,,$$

and, thus, $\widehat{u}(\eta) = Be^{-\eta}$ with an arbitrary constant $B \in \mathbb{R}$. Combining these results with the general solution (88) of the reduced equation gives

$$u(x) = Ae^{-x} + Be^{(x-1)/\varepsilon} + O(\varepsilon)\,,$$

and, with the boundary conditions in (86),

$$u(x) = e^{-x} - e^{-1+(x-1)/\varepsilon} + O(\varepsilon)\,.$$

This result can be easily verified by comparison with the exact solution of (86).

1.5 Regular/Singular perturbations of the oscillatory equation: the van der Pol equation and the "Method of Multiple Scales"

Let us consider the "oscillator(y) equation" for $u(t)$, where $\dot{u} = du/dt$:

$$\ddot{u} + u = 0\,, \tag{92}$$

with initial conditions

$$u(t=0) = \overline{u} \quad , \quad \dot{u}(t=0) = 0\,. \tag{93}$$

The solution is

$$u(t) = \overline{u}\,\cos(t)\,. \tag{94}$$

We now study 3 perturbations (i.e. $0 < \varepsilon \ll 1$) of (92), all with the same initial data (93).

perturbation-1 "simple case" :

$$\ddot{u} + u = \varepsilon u \tag{95}$$

This equation with initial data (93) has the exact solution

$$u_\varepsilon(t) = \overline{u}\,\cos\left((\sqrt{(1-\varepsilon)}\,t\right)\,. \tag{96}$$

The "reduced solution" for $\varepsilon = 0$, i.e. $u_0(t) = \overline{u}\,\cos(t)$, is a "formal approximation" for fixed time t, but not uniformly for all t since the small error in the frequency yields a large error for large t.
An improvement can be achieved by a Taylor expansion of the frequency in (96):

$$\sqrt{1-\varepsilon} = 1 - \frac{\varepsilon}{2} + O(\varepsilon^2)\,. \tag{97}$$

The resulting approximation

$$U_0(t) = \cos\left(t - \frac{\varepsilon t}{2}\right) \tag{98}$$

has much longer range of validity: while the reduced solution $u_0(t)$ approximates the exact solution as long as $t = o(\varepsilon^{-1})$, $U_0(t)$ is a valid approximation for $t = o(\varepsilon^{-2})$. However, the approximation with $U_0(t)$ is also not uniform in t.
By a further expansion of the frequency the range of validity can be further expanded, gaining "one order" $1/\varepsilon$ in time for one order ε in the expansion of the frequency.
In contrast, the standard "asymptotic expansion" of $u_\varepsilon(t)$ improves the accuracy on fixed bounded time intervals, but does not improve the range of validity.

perturbation-2 "friction case" :
Now add a "friction" term proportional to the "velocity" $\dot{u}$ to the unperturbed oscillator equation (92), (93):

$$\ddot{u} + u = -2\varepsilon\dot{u}\,. \tag{99}$$

Again we can compute the exact solution u_ε with initial data (93):

$$u_\varepsilon(t) = \overline{u}e^{-\varepsilon t}\cos\left(\sqrt{1-\varepsilon^2}\,t\right)\,. \tag{100}$$

Now not only the frequency, but also the amplitude of the oscillations is affected by the perturbation.

The long time behavior of the exact solution (100) and of the solution (94) of the reduced problem are significantly different, i.e. convergence to zero vs. periodic oscillations.
Again we can construct asymptotic expansions for fixed finite time t, but an approximation uniform in t, valid for arbitrary long times, cannot neglect the factor $e^{-\varepsilon t}$.
Both the perturbation-1 and the perturbation-2 are regular for fixed time t, but singular if we want an approximation for arbitrary long times $t \to \infty$.

perturbation-3: "van der Pol equation case" :
Let us consider another "more complicated" perturbation, the so called van der Pol equation. With initial data (93) it reads in scaled form:

$$\ddot{u} + u = \varepsilon(1 - u^2)\dot{u}\,. \tag{101}$$

Take the "reduced solution" from (94), i.e. $u_0(t) = \overline{u}\,\cos(t)$. Then the residual $u_\varepsilon - u_0$ given by

$$\varepsilon(1 - u_0^2)\dot{u}_0 = -\varepsilon(1 - \overline{u}^2\cos^2 t)\overline{u}\sin t \tag{102}$$

is $O(\varepsilon)$ uniformly in time, and therefore u_0 given by (94) is a "formal approximation" for the solution of (101), (93).
With the standard ansatz of an asymptotic expansion:

$$u = u_0 + \varepsilon u_1 + O(\varepsilon^2) \tag{103}$$

the correction term u_1 has to solve the problem

$$\ddot{u}_1 + u_1 = \overline{u}\left(\frac{\overline{u}^2}{4} - 1\right)\sin t + \frac{\overline{u}^3}{4}\sin 3t\,, \qquad u_1(0) = \dot{u}_1(0) = 0\,. \tag{104}$$

The first term on the right hand side oscillates with the same frequency as the solutions of the homogeneous equation. This produces "resonance":

$$u_1(t) = \frac{\overline{u}}{2}\left(1 - \frac{\overline{u}^2}{4}\right)(t\,\cos t - \sin t) + \frac{\overline{u}^3}{32}(3\,\sin t - \sin 3t)\,, \tag{105}$$

i.e. the amplitude of the oscillations in u_1 grows linearly in t. This also produces a linearly growing residual, which converges to zero as $\varepsilon \to 0$ pointwise in t, but not uniformly. Also the "correction" εu_1 is not small compared to u_0 for large t, since we have the term εt.

The appearance of the term εt in the asymptotic approximations in the 3 examples might lead to the conclusion that a bad scaling of time has been chosen. However, while for the modification of the frequency and of the amplitude the rescaled time scale $T_1 = \varepsilon t$ is appropriate, the basic oscillations still happen on the original time scale t. We are in an intrinsic "2-scale situation".

The method of asymptotic expansions can be adapted to this situation with the *method of multiple scales*. It still relies on an expansion of the form

$$u = u_0 + \varepsilon u_1 + O(\varepsilon^2)\,, \tag{106}$$

but now with the "multiple scales ansatz":

$$u_k = u_k(T_0, T_1, T_2, \cdots) \tag{107}$$

for the coefficients, with the independent variables

$$T_i = \varepsilon^i t\,. \tag{108}$$

The chain rule implies

$$\dot{u}_k = \frac{\partial u_k}{\partial T_0} + \varepsilon \frac{\partial u_k}{\partial T_1} + O(\varepsilon^2)\,, \qquad \ddot{u}_k = \frac{\partial^2 u_k}{\partial T_0^2} + 2\varepsilon \frac{\partial^2 u_k}{\partial T_0 \partial T_1} + O(\varepsilon^2)\,. \tag{109}$$

At first glance this method of "multiple scales" seems a bad idea, since we approximate an ODE by a system of PDEs. However, it is easily seen that the coefficients can be determined recursively by solving ODEs; and one can see - not so easily - that the "multiple scales" save the ansatz (106).
Note that the boundary layer technique of the previous chapter actually is sort of a "multiple scale" technique, with a sum of 2 terms: $u(x,\xi) = u_0(x) + \tilde{u}_1(x/\varepsilon^\alpha)$. However, we now put 2 scales in the first order correction term u_1.

Let us perform the "method of multiple scales" for the perturbation-3, van der Pol equation: Plugging (107) into the equation (101) yields

$$\frac{\partial^2 u_0}{\partial T_0^2} + 2\varepsilon \frac{\partial^2 u_0}{\partial T_0 \partial T_1} + \varepsilon \frac{\partial^2 u_1}{\partial T_0^2} + u_0 + \varepsilon u_1 = \varepsilon (1 - u_0^2) \frac{\partial u_0}{\partial T_0} + O(\varepsilon^2)\,,$$

and comparing coefficients gives

$$\frac{\partial^2 u_0}{\partial T_0^2} + u_0 = 0\,, \tag{110}$$

$$\frac{\partial^2 u_1}{\partial T_0^2} + u_1 = -2 \frac{\partial^2 u_0}{\partial T_0 \partial T_1} + (1 - u_0^2) \frac{\partial u_0}{\partial T_0}\,. \tag{111}$$

Now we see how the coefficients can be determined recursively by solving ODEs.
The general solution of the equation for u_0 can be written as

$$u_0 = a(T_1, T_2, \cdots) \cos\left(T_0 + b(T_1, T_2, \cdots)\right)$$

with arbitrary a and b, independent from T_0. The initial conditions (93) imply

$$a(0, 0, \cdots) = \overline{u}\,, \qquad b(0, 0, \cdots) = 0\,.$$

The right hand side of the equation for u_1 is now given by

$$\begin{aligned} &2 \frac{\partial a}{\partial T_1} \sin\phi + 2a \frac{\partial b}{\partial T_1} \cos\phi - a \sin\phi \left(1 - a^2 \cos^2\phi\right) = \\ &= \left(2 \frac{\partial a}{\partial T_1} - a + \frac{a^3}{4}\right) \sin\phi + 2a \frac{\partial b}{\partial T_1} \cos\phi + \frac{a^3}{4} \sin 3\phi \end{aligned}$$

with $\phi = T_0 + b$. As in the direct expansion method above, the terms with $\sin\phi$ and $\cos\phi$ would produce resonance. But now we can use the remaining freedom in the choice of a and b to eliminate these *secular terms*.[1] This idea gives for a and b the ODEs

$$\frac{\partial a}{\partial T_1} = \frac{a}{8}\left(4 - a^2\right)\,, \qquad \frac{\partial b}{\partial T_1} = 0\,. \tag{112}$$

[1] This name originates in one of the early applications of the multiple scales method in problems of astronomy.

Looking for a formal approximation producing a $O(\varepsilon^2)$ residual, we may neglect the dependence of a and b on the variables T_2, $T_3, \ldots$, and in u_1 the dependence on T_1, $T_2, \ldots$. The equation for a has the asymptotically stable steady states $a = \pm 2$ and the unstable steady state $a = 0$. With the initial conditions for a and b we obtain

$$a = \frac{2\overline{u}}{\sqrt{\overline{u}^2 - (\overline{u}^2 - 4)e^{-\varepsilon t}}}\,, \qquad b = 0\,. \tag{113}$$

The "two scale reduced solution" $u_0(T_0, T_1) = u_0(t, \varepsilon t)$ that, in somewhat sloppy notation, still denote by $u_0(t)$

$$u_0(t) = \frac{2\overline{u}\cos t}{\sqrt{\overline{u}^2 - (\overline{u}^2 - 4)e^{-\varepsilon t}}} \tag{114}$$

is a uniform in t approximation at first order, i.e. $u_\varepsilon(t) = u_0(t) + O(\varepsilon)$. It approaches, as $t \to \infty$, a periodic oscillation with amplitude 2 and period 2π, i.e. it shows *limit cycle* behavior. The existence of such a limit cycle can be shown for the full problem.

Remark:
The van der Pol equation is a model for a simple electrical circuit consisting of a capacitance C, an inductance L, and a nonlinear resistor R with the voltage-current characteristic $U_R = g(I)$, where U_R is the voltage across R and I is the current. A linear g would correspond to Ohm's law for a linear resistor. Here we assume a non-monotone, nonlinear relation with $g(I) = -g_1 I + g_3 I^3$, $g_1, g_3 > 0$. The basic laws for capacitance and inductance give $I = C\frac{dU_C}{dt^*}$ and, respectively, $U_L = L\frac{dI}{dt^*}$, with time t^*. Finally, the sum of the voltages has to vanish: $U_R + U_C + U_L = 0$.
The time derivative of this equation (multiplied by C) can be written as a second order ODE for the current $I(t)$:

$$LC\frac{d^2 I}{dt^2} + I = (-g_1 C + 3g_3 C I^2)\frac{dI}{dt}\,, \tag{115}$$

called the *van der Pol equation.*
For a nondimensionalization we use as basic units Ampere (A), the SI unit for electric current, and seconds (s) for time. The units of the parameters are then obvious from the equation:

Quantity	SI unit
LC	s^2
$g_1 C$	s
$g_3 C$	$\mathrm{s\ A}^{-2}$

We introduce the dimensionless quantities

$$u := \sqrt{\frac{3g_3}{g_1}}\, I\,, \qquad t = \frac{t^*}{\sqrt{LC}}\,, \qquad \varepsilon = g_1\sqrt{\frac{C}{L}}\,, \tag{116}$$

which leads to the scaled version (101) of the van der Pol equation.
In the 1920s, the physicist Balthasar van der Pol observed oscillations in electrical circuits and found (t)his equation as a possible theoretical explanation,which can be achieved by studying the long time behavior.

2 Asymptotic analysis and PDE in hydrodynamics

2.1 Overview of model hierarchy

$$
\begin{array}{lcc}
\text{"All orders of } \varepsilon\text{":} & \text{"exact" equation} & \longleftrightarrow \quad \text{"exact" solution } u_\varepsilon \\
& \Big\downarrow \varepsilon\to 0 & \Big\downarrow \varepsilon\to 0 \text{ (weak/strong limit)} \\
O(\varepsilon), O(\varepsilon^2): & \text{approximate equations} & \longleftrightarrow \quad \text{approximate solutions } u_1, u_2, \dots \\
& \Big\downarrow \varepsilon=0 & \Big\downarrow \varepsilon=0 \text{ (weak/strong limit)} \\
O(1), \varepsilon = 0: & \text{reduced equation} & \longleftrightarrow \quad \text{reduced solution } u_0
\end{array}
$$

regular perturbations: Asymptotic expansion ("Hilbert" or "Chapman-Enskog expansion")
singular perturbations: Boundary layers, 2-scale expansions, etc.

1a) Burgers equation - the basic example:

$$
\begin{array}{lcl}
\text{viscous Burgers equation:} & \partial_t u + u\partial_x u = -\varepsilon \Delta u & (113) \\
& \Big\downarrow \varepsilon\to 0 & \\
\text{inviscid Burgers equation:} & \partial_t u + u\partial_x u = 0 & (114)
\end{array}
$$

For the viscous Burger equation, there exists a unique global classical solution, but the reduced problem, also called the Hopf equation, generally has (non unique) weak solutions only. After a finite time, they can develop discontinuities, e.g. a "shock"/"jump".

1) Fluid(Hydro)dynamics:

a) inviscid limit

$$
\begin{array}{cc}
\text{Navier-Stokes equation} & \varepsilon\Delta \mathbf{v} \\
\Big\downarrow \varepsilon\to 0 & \Big\downarrow \\
\text{Euler equation} & 0
\end{array}
$$

where ε is a "viscosity" constant. The unknowns in the Navier-Stokes and Euler equations are the density $\rho(x,t)$ and the velocity $\mathbf{v}(x,t)$ (a vector field$\in \mathbb{R}^3$). A priori, for the Euler equations classical solutions only exist for small times (until the first shock) and the concept of "viscosity solutions" can be used to define global (in time) unique weak solutions.

b) incompressible limit

$$
\begin{array}{c}
\text{compressible Navier-Stokes} \\
\Big\downarrow \varepsilon\to 0 \\
\text{incompressible Navier-Stokes or Euler}
\end{array}
$$

where ε is the "Mach number".

"Incompressible" implies the condition div$\mathbf{v} = 0$, which is what is left from the "continuity equation" as the PDE for mass conservation.

2 Asymptotic analysis and PDE in hydrodynamics

2.1 Overview of model hierachy

$$
\begin{array}{lcc}
\text{"All orders of } \varepsilon\text{":} & \text{"exact" equation} & \longleftrightarrow \quad \text{"exact" solution } u_\varepsilon \\
& \Big\downarrow \varepsilon\to 0 & \Big\downarrow \varepsilon\to 0 \text{ (weak/strong limit)} \\
O(\varepsilon), O(\varepsilon^2): & \text{approximate equations} & \longleftrightarrow \quad \text{approximate solutions } u_1, u_2, \dots \\
& \Big\downarrow \varepsilon=0 & \Big\downarrow \varepsilon=0 \text{ (weak/strong limit)} \\
O(1),\ \varepsilon = 0: & \text{reduced equation} & \longleftrightarrow \quad \text{reduced solution } u_0
\end{array}
$$

regular perturbations: Asymptotic expansion ("Hilbert" or "Chapman-Enskog expansion")
singular perturbations: Boundary layers, 2-scale expansions, etc.

1a) Burgers equation - the basic example:

$$
\begin{array}{ll}
\text{viscous Burgers equation:} & \partial_t u + u\partial_x u = -\varepsilon \Delta u \\
& \Big\downarrow \varepsilon\to 0 \\
\text{inviscid Burgers equation:} & \partial_t u + u\partial_x u = 0
\end{array}
\tag{117}
$$

For the viscous Burger equation there exists a unique global classical solution, but the reduced problem, also called the Hopf equation, generally has (non unique) weak solutions only. After a finite time, they can develop discontinuities, e.g. a "shock"/"jump".

1) Fluid(Hydro)dynamics:

a) inviscid limit

$$
\begin{array}{cc}
\text{Navier-Stokes equation} & \varepsilon\Delta \mathbf{v} \\
\Big\downarrow \varepsilon\to 0 & \Big\downarrow \\
\text{Euler equation} & 0
\end{array}
$$

where ε is a "viscosity" constant. The unknowns in the Navier-Stokes and Euler equations are the density $\rho(x,t)$ and the velocity $\mathbf{v}(x,t)$ (a vector field$\in \mathbb{R}^3$). A priori, for the Euler equations classical solutions only exist for small times (until the first shock) and the concept of viscosity solutions can be used to define global (in time) unique weak solutions.

b) incompressible limit

$$
\begin{array}{c}
\text{compressible Navier-Stokes} \\
\Big\downarrow \varepsilon\to 0 \\
\text{incompressible Navier-Stokes or Euler}
\end{array}
$$

where ε is the "Mach number". "Incompressible" implies div$\mathbf{v} = 0$, which replaces the equation for mass conservation.

2) From kinetic to hydrodynamic models:

$$\begin{array}{c} \text{Kinetic equation (Vlasov, Boltzmann)} \\ \Big\downarrow \varepsilon \to 0 \\ \text{(incompressible) Fluid equation (Navier-Stokes, Euler)} \end{array}$$

where ε is e.g. the "mean free path" between two particles/collisions. The unknown in the Boltzmann equation is a probability density $f(x, \mathbf{v}, t) \geq 0$ in "phase space" $\mathbb{R}^3_x \times \mathbb{R}^3_{\mathbf{v}}$, which is a "microscopic" view. In the "hydrodynamic" equations, the unknowns are densities in "physical space", like the mass density $\rho(x, t)$ and velocity/momentum $\mathbf{v}(x, t)$, which are "macroscopic".

The above 2 diagrams sketch a huge program of "asymptotic analysis of nonlinear PDEs", with good progress by many great mathematicians in the last 30 years. We just mention a few, with close ties to the Vienna Applied PDE school (founded by Peter A. Markowich), who still give key contributions, in 3 generations: The French Applied PDE schools of Pierre-Louis Lions and his students Benoit Perthame, Nader Masmoudi, Cedric Villani,... and the school of Claude Bardos and his students Pierre Degond, Francois Golse and his student Laure St. Raymond, and Dave Levermore, a student of Peter Lax. Of course, one should mention Jean Leray, one of the founding fathers - indeed "Leray solutions" are a basic concept of (weak) solutions in the field.

3) From point-particle mechanics to fluid mechanics:

$$\begin{array}{c} \text{point-particle mechanics } (N \text{ particles}) \\ \Big\downarrow \varepsilon = \frac{1}{N} \to 0 \\ \text{continuum mechanics} \end{array}$$

On the most fundamental "atomistic" - "microscopic" level, in point-particle classical mechanics, the unknowns are the positions $x_j(t)$ and velocities $v_j(t)$, and with Newton's 2nd law, $F = ma$, we get N 2nd order ODEs or $2N$ 1st order ODEs for N particles.

In continuum mechanics, there is a scalar PDE for the density $\rho(x, t)$, a vector equation for the 3 components of the velocity ("momentum", "impulsion",...) $\mathbf{v}(x, t)$, and another scalar PDE for energy $e(x, t)$ and temperature $T(x, t)$. These five nonlinear PDEs are the "fundamental equations of fluid mechanics" that we shall present in more detail, with the (Navier) Stokes and Euler equations as special cases.

Note that the above limits correspond to problem 6b of Hilbert's famous 23 problems presented at ICM in Paris on 8 Aug 1900: *"The rigorous theory of limiting processes which lead from the atomistic view to the laws of motion of continua."*
Note further that the open problem of global in time classical (?) unique (?) solutions of the Euler / Navier Stokes equations is one of the 7 "millenium problems" where the Clay Institute would pay 1 Million USD for a solution.

Let us mention that there are many more such model hierarchies in natural sciences, e.g. the hierarchy from quantum mechanics to classical mechanics, the "semi-classical limit" when the Planck constant vanishes $\varepsilon = \hbar \to 0$, from the Schrödinger to the Vlasov equation, or the hierarchy of relativistic (Einstein) to nonrelativistic (Newton, Schrödinger) physics when the speed light c goes to infinity $\varepsilon = 1/c \to 0$.

2.2 Point particle mechanics: some concepts

Consider N mass points interacting by pairwise central forces and under the action of additional external forces. The position of the ith particle at time t and its mass are denoted by

$$r_i(t) \in \mathbb{R}^3 \quad \text{and} \quad m_i\,, \qquad i = 1, \ldots, N\,. \tag{118}$$

For the momenta $p_i = m_i \dot{r}_i$ ($\dot{r} = dr/dt$), Newton's second law implies

$$\dot{p}_i = \sum_{j \neq i} f_{ij} + f_i\,, \qquad i = 1, \ldots, N\,, \tag{119}$$

where f_{ij} is the force exerted by the jth particle on the ith particle, and f_i is the external force acting on the ith particle. Newton's third law $f_{ij} = -f_{ji}$ is satisfied by central forces of the form

$$f_{ij}(r_i, r_j) = \frac{r_i - r_j}{|r_i - r_j|} g_{ij}(|r_i - r_j|)\,,$$

with $g_{ij} = g_{ji}$. For $g_{ij} > 0$ the interaction is repulsive, for $g_{ij} < 0$ it is attractive. The external forces $f_i(t)$ are assumed as given. Two examples:

1. Gravitation:
$$g_{ij}(r) = -G\frac{m_i m_j}{r^2} \tag{120}$$
with the gravitational constant $G > 0$,

2. Electrostatics:
$$g_{ij}(r) = k\frac{Q_i Q_j}{r^2} \tag{121}$$
with the charge Q_i of the ith particle and the proportionality constant $k > 0$. Since charges can be positive or negative, the force can be repulsive or attractive.

With the total momentum $p := \sum_i p_i$ and the sum of the external forces $f := \sum_i f_i$ we obtain

$$\dot{p} = \sum_i \sum_{j \neq i} f_{ij} + f = \sum_i \left(\sum_{j<i} f_{ij} + \sum_{j<i} f_{ji} \right) + f = f\,, \tag{122}$$

because of Newton's third law. This means that the total momentum is not influenced by the interaction forces. This equation can be interpreted as the equation for one point particle, where the total mass $m = \sum_i m_i$, is concentrated in the center of mass $r = (1/m)\sum_i m_i r_i$, $p = m\dot{r}$ is the total momentum, and its motion is determined by the total external force. In the absence of external forces we talk about a *closed system*, exhibiting *conservation of total momentum*, i.e. $p(t) = \text{const}$. In this case the movement of the center of mass is inertial motion with constant velocity.

An analogous result holds for *angular momentum*. The angular momentum with respect to the position r_0 of the ith particle is given by

$$L_i = (r_i - r_0) \times p_i \tag{123}$$

with the vector product $\times$, and the *torque* (time derivative of angular momentum, corresponding to 'force' for the linear momentum) exerted by the ith external force is

$$M_i = (r_i - r_0) \times f_i\,, \qquad i = 1, \ldots, N\,. \tag{124}$$

The total angular momentum and the total torque due to external forces are given by $L = \sum_i L_i$ and, respectively, $M = \sum_i M_i$. By the property $a \times a = 0$ of the vector product, $\dot{L}_i = (r_i - r_0) \times \dot{p}_i$ holds and, thus,

$$\begin{aligned} \dot{L} &= \sum_i (r_i - r_0) \times \dot{p}_i = \sum_i \sum_{j \neq i} r_i \times f_{ij} - r_0 \times \sum_i \sum_{j \neq i} f_{ij} + M \\ &= -\sum_i \sum_{j \neq i} \frac{r_i \times r_j}{|r_i - r_j|} g_{ij}(|r_i - r_j|) + M = M \,, \end{aligned} \tag{125}$$

by the antisymmetry of the vector product. Again the interaction forces do not contribute to the total torque acting on the system. In a closed system ($f = 0$) we have *conservation of total angular momentum*, i.e. $L(t) = \text{const}$.

Finally we investigate the change of the total *kinetic energy* T

$$T = \sum_i \frac{m_i |\dot{r}_i|^2}{2} \,, \tag{126}$$

where we need the total *work* W done by the external forces:

$$W = \sum_i \dot{r}_i \cdot f_i \,. \tag{127}$$

We compute

$$\begin{aligned} \dot{T} &= \sum_i \dot{r}_i \cdot \dot{p}_i = \sum_i \sum_{j \neq i} \dot{r}_i \cdot f_{ij} + W \\ &= \sum_i \sum_{j<i} \frac{(\dot{r}_i - \dot{r}_j) \cdot (r_i - r_j)}{|r_i - r_j|} g_{ij}(|r_i - r_j|) + W \\ &= \sum_i \sum_{j<i} g_{ij}(|r_i - r_j|) \frac{d}{dt} |r_i - r_j| + W \,. \end{aligned}$$

With the primitives G_{ij} of g_{ij}, $j < i$, we define the potential energy

$$V = -\sum_i \sum_{j<i} G_{ij}(|r_i - r_j|) \,, \tag{128}$$

and the total energy $E = T + V$, satisfying

$$\dot{E} = W \,. \tag{129}$$

As for (angular) momentum, the interaction forces do not contribute to the energy balance, and for a closed system ($f_i = 0$, $i = 1, \ldots, N \quad \Rightarrow \quad W = 0$) we have *conservation of energy*.

The 1+3+1 conservation laws of mass, (angular) momentum and energy for point particle mechanics can be converted into 1+3+1 PDE for continuum/fluid mechanics. We shall introduce and discuss this huge class of coupled nonlinear PDE systems in the sequel.

Gravitation and electric forces lead to permanent interactions between all particles. The above conclusions are also valid for other systems.

An instructive example is the ***hard sphere model*** for rarefied gases:

We consider an ensemble of N spherical particles with mass m and radius R, moving in $\mathbb{R}^3$. Let $r_i(t)$ denote the center of the ith sphere at time t, $i = 1, \dots, N$. As long as the particles do not collide, i.e. $|r_i - r_j| > 2R$, $\forall\, i \neq j$, the motion of the particles is influenced only by the external forces. Now consider a collision between the ith and the jth particle at time t_0, i.e.

$$|r_i(t_0) - r_j(t_0)| = 2R\,.$$

We assume the collision to be *hard*, which means that they lead to an instantaneous change of the velocities of the two particles. The assumption of central forces then means that the change of velocities is in the direction of the vector

$$n = \frac{r_i(t_0) - r_j(t_0)}{2R}\,.$$

Denoting the precollisional velocities by $v_i = \dot{r}_i(t_0-)$ and the postcollisional velocities by $v_i' = \dot{r}_i(t_0+)$ and taking into account Newton's third law, we have

$$v_i' = v_i + \lambda n\,, \qquad v_j' = v_j - \lambda n\,.$$

The constant λ is determined by the assumption that the collision is *elastic*, meaning that the total kinetic energy is conserved:

$$|v_i'|^2 + |v_j'|^2 = |v_i|^2 + |v_j|^2\,.$$

This gives

$$\lambda = n \cdot (v_j - v_i)\,.$$

It is easily checked that we also have conservation of momentum and of angular momentum:

$$v_i' + v_j' = v_i + v_j\,, \qquad r_i \times v_i' + r_j \times v_j' = r_i \times v_i + r_j \times v_j\,.$$

Thus, also this "3-D billiards" system satisfies the balance laws (122), (125), and (128), where the total energy between collisions is only kinetic energy.

For more realistic, continous, "binary" interaction between "atoms/molecules" of a rarified gas, L. Boltzmann introduced the "kinetic" PDE that bears his name. The unknown $f(x, v, t) \geq 0$ is a nonnegative "probability distribution" to find a particle at position x with "velocity" v at time t. In the **Boltzmann equation** for elastic collisions, the conservation of mass, momentum and energy hold, but a quantitity called "entropy" (essentially given by $f\ log(f)$) can only increase in the time evolution, as expressed in the famous "H theorem". The entropy is closely related to information theory and measures "(dis)order" and is linked to "irreversibility". Without stating the equation we just mention that the theory of the Boltzmann equation, e.g. existence of global solutions, yielded two Fields medals (P.-L. Lions and C. Villani).
The collisionless Boltzmann equation, the **Vlasov equation** is also a rich source of mathematical problems in modelling, analysis and numerics. These kinetic equations in nonlinear form (e.g. coupled to the Poisson or Maxwell equations) are very important e.g. in plasma physics, astrophysics and also biology etc.

2.3 From point particles to a continuous medium

The transition from the discrete N particle system to a continous medium can be performed in the limit $N \to \infty$ while keeping the "density" of particles finite.

We denote the cube with center $\mathbf{x}_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ and edge length $h > 0$ by

$$W_h(\mathbf{x}_0) = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; \max\left(|x - x_0|, |y - y_0|, |z - z_0|\right) \leq \frac{h}{2} \right\}.$$

Consider a collection of $N \gg 1$ point masses, e.g. molecules like in the previous subsection. It is possible to define a mass density ρ_h by

$$\rho_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \,. \tag{130}$$

This density should have the property that the integral

$$\int_R \rho_h(\mathbf{x}, t)\, d\mathbf{x} \tag{131}$$

can be used as an approximation for the mass contained in $R \subset \mathbb{R}^3$. To achieve this, one has to choose a small h in comparison to the characteristic length R. On the other hand if h is too small such that there aren't many molecules in the cube $W_h(\mathbf{x}_0)$ then the defined density function is not very meaningful because it is very dependent on h. The basic assumption of continuum mechanics is that there exists a range of values of h for which ρ_h is only weakly dependent on h. For the following we will assume that h lies in this range and that the integral (130) is the exact formula for the mass contained in R. Similarly to (129) we are able to define the momentum density

$$\mathbf{p}_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \dot{\mathbf{r}}_i(t) \,, \tag{132}$$

and with that a mean velocity $\mathbf{v}_h = \frac{\mathbf{p}_h}{\rho_h}$. Furthermore the energy density $E_h = T_h + V_h$ is given by the sum of the density of the kinetic energy

$$T_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \frac{|\dot{\mathbf{r}}_i(t)|^2}{2}$$

and the similarly defined density of the potential energy. If we define a kinetic energy contribution by

$$\widetilde{T}_h(\mathbf{x}_0, t) = \frac{1}{h^3} \sum_{\mathbf{r}_i(t) \in W_h(\mathbf{x}_0)} m_i \frac{|\dot{\mathbf{r}}_i(t) - \mathbf{v}_h|^2}{2} \,,$$

that is caused by a deviation of the particle velocities to the mean velocity then we obtain

$$T_h = \rho_h \frac{|\mathbf{v}_h|^2}{2} + \widetilde{T}_h \,.$$

Finally we introduce the microscopic or internal energy per mass unit e_h by the equation $\rho_h e_h = \widetilde{T}_h + V_h$. Using this, we can write the energy density as sum of the density of the macroscopic kinetic energy and the internal energy:

$$E_h = \rho_h \left(\frac{|\mathbf{v}_h|^2}{2} + e_h \right).$$

Note that this approach to continuum mechanics relies on averaging as in (129) and (131).

The system of $2N$ coupled (by the interaction) ODEs for the trajectories of N point particles, becomes a system of 5 coupled nonlinear PDEs for the mass and velocity densities $\rho(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, t)$ and an energy density $e(\mathbf{x}, t)$.

Remark: If one wants to avoid the "atomistic" picture, one can start from the picture that mass is distributed continuously in space and introduce the existence of a mass density $\rho(\mathbf{x}, t)$ and an average velocity $\mathbf{v}(\mathbf{x}, t)$, without making any considerations about the microscopic structure of the material studied.

2.3.1 Euler vs Lagrangian coordinates - material derivative

In the study of the kinematics of continuous mediums it is useful to work with two different settings of "coordinates" and with the notion of a "material derivative".

If the average velocity $\mathbf{v}(\mathbf{x}, t)$ is given, the motion of a material particle can be described by the solution of the initial value problem

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, t) \\ \mathbf{x}(0) = \mathbf{A} \end{cases},$$

where $\mathbf{A}$ denotes the position of the particle at time $t = 0$. We introduce the dependence of the particle trajectory on the initial position by the equation

$$\mathbf{x} = \mathbf{x}(\mathbf{A}, t)\,. \tag{133}$$

Therefore $\mathbf{x}(\mathbf{A}, t)$ is the position of the particle at time t, which had the initial position $\mathbf{A}$ at time 0. We assume impenetrability of the material. This means that equation (132) can be uniquely solved for $\mathbf{A}$ and is equivalent to an equation of the form

$$\mathbf{A} = \mathbf{A}(\mathbf{x}, t). \tag{134}$$

Consequently every function of $(\mathbf{x}, t)$ can be expressed as a function of $(\mathbf{A}, t)$. In particular we introduce for the mass density $\rho(\mathbf{x}, t)$ of a particle at time t, which had position $\mathbf{A}$ at time 0 the notation

$$\delta(\mathbf{A}, t) = \rho\left(\mathbf{x}(\mathbf{A}, t), t\right).$$

Depending on whether we use *Euler-coordinates* $\mathbf{x}$ or *Lagrange-coordinates* $\mathbf{A}$, we call this a *position-dependent* or *material-dependent* description.

We observe the following relation between the partial derivative of time with fixed Euler-coordinates and with fixed Lagrange coordinates. It holds that

$$\frac{\partial \delta}{\partial t} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho\,,$$

where the gradient $\nabla\rho$ of ρ is the vector of the partial derivative of ρ of the components of the position vector $\mathbf{x}$. We call the right hand side the *material derivative* (aka "mass derivative") of ρ and introduce the notation

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho. \tag{135}$$

2.4 The "constitutive equations" of continuum mechanics

The following system of 1+3+1 nonlinear coupled PDEs are the "fundamental equations" of "continuum mechanics" where a system (of N "particles") is described by continuous densities, in particular the scalar "mass density" $\rho(x,t)$ and the vector "velocity (or "momentum") field" $\mathbf{v}(x,t)$:

- Conservation of mass:
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{136}$$

- Conservation of momentum:
$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{T}) = \rho \mathbf{f} \tag{137}$$
Here $\mathbf{x} \otimes \mathbf{y} = \mathbf{x}\mathbf{y}^{tr}$ is the tensor product of two vectors (aka "dyadic product"), namely a matrix with entries $x_i y_j$, $i,j = 1,2,3$. The term $\rho \mathbf{v} \otimes \mathbf{v}$ can be interpreted as **momentum flux**. The tensor $\mathbf{T}$ is called "stress tensor", $\mathbf{f}$ is an external force.

- Conservation of angular momentum:
$$\mathbf{T} = \mathbf{T}^{tr} \tag{138}$$

- Conservation of energy:
$$\frac{\partial}{\partial t}\left[\rho\left(\frac{|\mathbf{v}|^2}{2} + e\right)\right] = \rho \mathbf{f} \cdot \mathbf{v} - + \nabla \cdot \left[\rho \mathbf{v}\left(\frac{|\mathbf{v}|^2}{2} + e\right) + \mathbf{q} - \mathbf{T} \cdot \mathbf{v}\right] \tag{139}$$

The unknowns in these equations are the density ρ, the velocity $\mathbf{v}$, the (Cauchy) stress tensor $\mathbf{T}$, the specific internal energy e and the heat flux density $\mathbf{q}$. We observe that these are 17 unknown scalar functions. The conservation of angular momentum, i.e. the symmetry of the stress tensor, eliminates 3 unknowns. We have 5 differential equations at our disposal. Combining these shows that the system (135)–(138) is strongly underdetermined. This is expected, since these equations should describe a large class of materials. At this point, we have not used any information about the properties of the material, for example whether it is a gas, fluid or solid.

To create a determined system of equations we need additional equations, so called *constitutive equations*. Some fundamental examples are the following:

- In a rigid body the distance between two particles is constant.

- For a frictionless fluid or gas the (Cauchy) "stress vector" is given by $\mathbf{T} = -p\mathbf{n}$, where n denotes the normal vector and p denotes the pressure.

- Newton-Fourier's law of cooling states, that the heat flux density is proportional to the gradient of the temperature:
$$\mathbf{q} = -\kappa \nabla T$$
where T denotes the temperature and κ the thermal conductivity.

- For an ideal gas the relation $p = \rho RT$ holds, where R is a constant dependent on the gas. Additionally we often make the assumption that the internal energy is proportional to the change of temperature: $e = c_v T + \text{const}$, where c_v is the specific heat for constant volume.

The origins of constitutive equations are diverse. Some are just heuristic simplifications of real physics, sometimes they can be verified by experiments. Only for few cases a deduction from microscopic behavior is known.

2.4.1 Constitutive equations using material derivatives

Using the notation of material (mass) derivatives (134), the above PDE (135),(136) and(138) of subsection 2.4 can be rewritten as

- Conservation of mass:
$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \tag{140}$$
- Conservation of momentum:
$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{f} + \nabla \cdot \mathbf{T} \tag{141}$$
- Conservation of energy:
$$\rho \frac{D}{Dt} \left(\frac{|\mathbf{v}|^2}{2} + e \right) = \rho \mathbf{f} \cdot \mathbf{v} - \nabla \cdot (\mathbf{q} - \mathbf{T} \cdot \mathbf{v}) \tag{142}$$

For the case of vicuous fluids, the stress tensor $\mathbf{T}$ has one component given by the (scalar) pressure p, and another component caused by viscosity depending on the derivative of the velcity field $\mathbf{v}$. Symmetry considerations lead to the relation $\mathbf{T} = -p\mathbf{I} + \lambda(\nabla \cdot \mathbf{v})\, \mathbf{I} + \mu(\nabla \otimes \mathbf{v} + (\nabla \otimes \mathbf{v})^{tr})$, where $p = p(x,t)$ is the pressure, $\mathbf{I}$ the identity matrix, μ the shear viscosity , and λ another constant describing viscosity effects.

2.5 Fluid dynamics: Navier Stokes and Euler equations (in $\mathbb{R}^3$, incompressible case)

A common assumption which is used for typical fluids like water is incompressibility:

$$\frac{D\rho}{Dt} = 0.$$

An even stronger assumption is that the density is constant $\rho \equiv \text{const} = \rho_0$.

For this situation, the "fundamental equations" take the following form: the PDE (136) for the conservation of mass becomes the "incompressibility condition":

$$div\, \mathbf{v} \;=\; \nabla \cdot \mathbf{v} = 0 \tag{143}$$

plus the 3 PDEs (137) from the conservation of momentum

$$\frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\vec{\nabla}\mathbf{v} = \mathbf{f} - \nabla p + \mu \Delta \mathbf{v} \tag{144}$$

These are the incompressible Navier-Stokes equations (written with material derivative), that form a closed system of 1+3 PDEs (one scalar equation for the incompressibility, and 3 components of the vector equation for the momentum) for the 4 unknowns: the pressure p, and the 3 components of the velocity $\mathbf{v}$. Note that in (144) the gradient of p is a vector, and $\Delta \mathbf{v} = \frac{\partial^2 \mathbf{v}}{\partial x^2} + \frac{\partial^2 \mathbf{v}}{\partial y^2} + \frac{\partial^2 \mathbf{v}}{\partial z^2}$, and the second term in the material derivative of $\mathbf{v}$ is: $\mathbf{v} \cdot \vec{\nabla}\mathbf{v} = \vec{\nabla}(\mathbf{v} \otimes \mathbf{v}) = \mathrm{v}_x \cdot \frac{\partial \mathbf{v}}{\partial x} + \mathrm{v}_y \cdot \frac{\partial \mathbf{v}}{\partial y} + \mathrm{v}_z \cdot \frac{\partial \mathbf{v}}{\partial z}$.

For the non-viscous case, i.e. when the viscosity coefficient μ vanishes in (144), we get the incompressible Euler equation:

$$\frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\vec{\nabla}\mathbf{v} = \mathbf{f} - \nabla p \tag{145}$$

In the framework of Section 2.1, the non-viscous limit $\mu \to 0$ means to pass from a second order parabolic equation with smooth solutions to a first order, nonlinear hyperbolic equation, in which "shocks" can appear in finite time.

Note that in the incompressible setting many questions are still open: the presence of the pressure gradient term (which acts as a nonlocal operator) significantly changes the situation, in particular in comparison to the simpler situation of the Burgers equation (117) and other textbook quasilinear parabolic and hyperbolic systems. In two space dimensions, it is in fact known that both Navier-Stokes and Euler have global smooth solutions, while no proof (finite-time blow up or global well-posedness) exists for either in 3D. This ("Millenium") problem of "existence and smoothness of Navier Stokes equations" was laid out, e.g., in a concise paper of C. Fefferman.

2.5.1 Boundary layers in fluid dynamics - a model problem

As a general rule, much can be learned by investigating the *steady states* of a physical system, which constitute the time-independent solutions of the governing equations. In the context of incompressible fluids, this leads to a stationary version of 143 and 144, known as the Stokes equations:

$$(\mathbf{v}\cdot\vec{\nabla})\mathbf{v} = \mathbf{f} - \frac{1}{\rho_0}\nabla p + \mu\Delta\mathbf{v}, \tag{146a}$$

$$\nabla\cdot\mathbf{v} = 0. \tag{146b}$$

Looking at the scaling of the Stokes equations, we observe that the flow is only influenced by the pressure gradient, but not the level of pressure; therefore, we do not include a reference pressure in our nondimensional parameters. Choosing appropriate length and velocity scales L and U (which depend on the problem under consideration), we then obtain the Stokes equations in nondimensional form:

$$(\mathbf{v}\cdot\vec{\nabla})\mathbf{v} = \mathbf{f} - \nabla p + \frac{1}{Re}\Delta\mathbf{v}, \tag{147a}$$

$$\nabla\cdot\mathbf{v} = 0, \tag{147b}$$

where we introduced the Reynolds number $Re = \frac{UL}{\nu}$, with the kinematic viscosity $\nu = \frac{\mu}{\rho_0}$.
Now, when studying slightly viscous flow in a region with a solid boundary such as a channel (or the surroundings of a body embedded in the fluid), Re will be very large away from the boundary. The flow will therefore be effectively inviscid everywhere except in a layer around the boundary, where viscosity cannot be neglected. Thus, setting

$$\epsilon := \frac{1}{Re},$$

we arrive at a singular perturbation problem. In general, this problem does not permit an explicit solution; however, in certain domains, (147) reduces to a system that is analytically tractable. A particularly convenient example is provided by radial flow in a sector with a sink at the origin, treated comprehensively in the supplementary notes.

Remarks on the supplementary notes:

- The flow is directed inward, that is, f in (3.2.19) is nonpositive.
- The "dimensional reasoning" invoked to derive (3.2.19) and (3.2.20) relies on a similarity argument. (*Exercise*)
- The matching conditions, which are here derived by introduction of a matching variable, can also be found by requiring that the boundary layer correction terms vanish as $\theta^\star \to \infty$.