CHAPTER 1

Introduction

Throughout the book the usual Einstein summation convention is used: whenever an index appears twice in one expression, the summation over that index is performed. The term **domain** is reserved for an open set in a euclidean space.

1.1 Examples of evolution systems

Most of the equations and systems studied in this book have physical origin. They can be derived from basic balance laws by taking particular forms for constitutive relations and/or by considering particular kind of materials and processes.

Let $\Omega = \bigcup_{t=0}^{T} \Omega_t$, where $\Omega_t \subset \mathbb{R}^d$ is a domain occupied by the material at an instant of time $t \in [0, T)$, $T > 0$. Then the following system of equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho v_j)}{\partial x_j} = 0\,, \tag{1.1}$$

$$\frac{\partial (\rho v_i)}{\partial t} + \frac{\partial (\rho v_i v_j)}{\partial x_j} = \frac{\partial T_{ij}}{\partial x_j} + \rho f_i\,, \quad 1 \le i \le d\,, \tag{1.2}$$

$$\frac{\partial (\rho e)}{\partial t} + \frac{\partial (\rho e v_j)}{\partial x_j} = \frac{\partial (T_{kj} v_k)}{\partial x_j} - \frac{\partial q_j}{\partial x_j} + \rho r + \rho f_j v_j \tag{1.3}$$

represents the local forms of the law of conservation of mass, the law of balance of momentum and balance of energy, respectively. See for example TRUESDELL [1991] or CHADWICK [1976]. Here, ρ is the density, $\mathbf{v}$ is the velocity field, E is the specific internal energy of the material, $\mathbf{T}$ is the symmetric stress tensor, $\mathbf{q}$ is the spatial heat flux vector, r is the rate of external communication of heat to the body through radiation, $\mathbf{f}$ represents the specific external body forces and e denotes $E + |\mathbf{v}|^2/2$. All quantities are evaluated at $(t, x) \in [0, T) \times \Omega_t$. In the sequel we consider $\Omega_t = \Omega$ for all $t \ge 0$.

We note that (1.1)–(1.3) can have various forms, for example when the notion of material derivative is used. Let us recall that the **material derivative** of a scalar- (or vector- or tensor-) valued function $\xi : (0,T) \times \Omega \longrightarrow \mathbb{R}$ (or $\mathbb{R}^d$ or $\mathbb{R}^{d^2}$, respectively) is defined by

$$\frac{d\xi}{dt} \equiv \frac{\partial \xi}{\partial t} + v_j \frac{\partial \xi}{\partial x_j} .$$

In view of this, (1.1)–(1.3) can be rewritten as

$$\frac{d\rho}{dt} + \rho \frac{\partial v_j}{\partial x_j} = 0 , \tag{1.4}$$

$$\frac{d(\rho v_i)}{dt} + \rho v_i \frac{\partial v_j}{\partial x_j} = \frac{\partial T_{ij}}{\partial x_j} + \rho f_i , \quad 1 \le i \le d , \tag{1.5}$$

$$\frac{d(\rho e)}{dt} + \rho e \frac{\partial v_j}{\partial x_j} = \frac{\partial (T_{kj} v_k)}{\partial x_j} - \frac{\partial q_j}{\partial x_j} + \rho r + \rho f_j v_j . \tag{1.6}$$

Now, we define hyperbolic systems of partial differential equations, an important example of a large class of evolution systems.

Definition 1.7 *Let $\mathcal{O} \subseteq \mathbb{R}^s$ be a domain and let $\mathbf{f}_j, 1 \le j \le d$, be smooth functions from $\mathcal{O}$ into $\mathbb{R}^s$. The system of partial differential equations*

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}_j(\mathbf{u})}{\partial x_j} = \mathbf{0} \qquad in \quad \mathbb{R}^+ \times \mathbb{R}^d$$

is said to be **hyperbolic** *if for every $\mathbf{u} \in \mathcal{O}$ and every vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{R}$, the matrix*

$$\mathbf{J}(\mathbf{u}, \boldsymbol{\alpha}) \equiv \alpha_j \, \mathbf{J}_j(\mathbf{u})$$

has s real eigenvalues and is diagonalizable. Here we denote by $\mathbf{J}_j(\mathbf{u})$ the Jacobian matrix of the function $\mathbf{f}_j(\mathbf{u})$, $j = 1, \dots, d$, with the components

$$[\mathbf{J}_j(\mathbf{u})]_{ik} = \frac{\partial f_{ji}}{\partial u_k}(\mathbf{u}) , \qquad 1 \le i, k \le s .$$

1.1.1 Euler equations in 2D

Let us consider an inviscid (ideal) non-heat-conductive fluid for which both heat sources and external forces are neglected, i.e.,

$$\begin{gathered}\mathbf{T} = -p\,\mathbf{I}\,,\\ \frac{\partial q_j}{\partial x_j} = 0\,, \quad r = 0\,, \quad \mathbf{f} = \mathbf{0}\,,\end{gathered} \tag{1.8}$$

respectively.

Under the assumptions (1.8) the right-hand side of (1.2) is equal to $-\frac{\partial p}{\partial x_i}$, while (1.3) takes the form of

$$\frac{\partial(\rho e)}{\partial t} + \frac{\partial(\rho e v_j)}{\partial x_j} = -\frac{\partial(p v_j)}{\partial x_j}\,. \tag{1.9}$$

In such a way the system (1.1)–(1.3) turns into the so-called **Euler system** $(1 \leq i \leq d)$

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v_j)}{\partial x_j} = 0\,, \tag{1.10}$$

$$\frac{\partial(\rho v_i)}{\partial t} + \frac{\partial(\rho v_i v_j + p\delta_{ij})}{\partial x_j} = 0\,, \tag{1.11}$$

$$\frac{\partial(\rho e)}{\partial t} + \frac{\partial\left((\rho e + p)v_j\right)}{\partial x_j} = 0\,, \tag{1.12}$$

where δ_{ij} stands for the Kronecker delta. As before, all the functions are defined on $[0,T) \times \Omega$; $T \in (0,\infty]$, $\Omega \subseteq \mathbb{R}^d$ being a domain.

According to classical literature (Pippard [1957], Sommerfeld [1964], Courant and Friedrichs [1948]) only two of the quantities describing the thermodynamic state of a fluid (as the pressure p, the temperature θ, the density ρ, the internal energy E, the entropy η) are independent. Considering ρ, θ to be independent, one obtains

$$p = p(\rho,\theta)\,, \tag{1.13}$$

$$E = E(\rho,\theta)\,, \tag{1.14}$$

$$\eta = \eta(\rho,\theta)\,, \tag{1.15}$$

and

$$e = e(\rho,\theta,\mathbf{v})\,. \tag{1.16}$$

Therefore, in general, the Euler system (1.10)–(1.12) together with

(1.13)–(1.16) is considered as the system of $(d+2)$ equations for $(d+2)$ unknown functions ρ, θ, v_j. However, the knowledge of explicit relations of the type (1.13), (1.14) allows us to consider the Euler system as a system for another set of unknown functions.

As an example, let us consider the state equation of a **perfect** gas

$$p = R\rho\theta\,, \tag{1.17}$$

where R is the universal gas constant. Further, let us consider a **polytropic** perfect gas, for which we have

$$E = c_V\theta\,, \tag{1.18}$$

where the constant c_V is the specific heat at constant volume. Then we have $p = \frac{R}{c_V}\rho E$ and, setting $\gamma \equiv 1 + R/c_V > 1$ (usually called the **Poisson constant**), we arrive at

$$p = (\gamma - 1)\rho\left(e - \frac{|\mathbf{v}|^2}{2}\right)\,. \tag{1.19}$$

Hence, in this situation, the Euler system (1.10)–(1.12) together with the state equation (1.19) can be viewed as a system of $(d+2)$ equations for $(d+2)$ unknown functions ρ, e, v_j.

If we now set

$$\mathbf{u} = (\rho, \rho v_1, \ldots, \rho v_d, \rho e)^T \in \mathbb{R}^s\,, \quad s = d+2\,,$$

we can rewrite (1.10)–(1.12) (under an additional assumption that $\rho > 0$) as a quasilinear system of partial differential equations

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}_j(\mathbf{u})}{\partial x_j} = \mathbf{0}\,, \tag{1.20}$$

where

$$\mathbf{f}_j(\mathbf{u}) \equiv \left(\rho v_j, \rho v_1 v_j + \delta_{1j}p, \ldots, \rho v_d v_j + \delta_{dj}p, (\rho e + p)\, v_j\right)^T\,,$$

$j = 1, \ldots, d$ and $\mathbf{f}_j \in C^1(\mathcal{O}; \mathbb{R}^s)$.

In particular, for $d = 2$, we have $\mathbf{u} = \left(\rho, \rho v_1, \rho v_2, \rho e\right)^T$ and

$$\mathbf{f}_1(\mathbf{u}) = \begin{pmatrix} \rho v_1 \\ \rho v_1^2 + p \\ \rho v_1 v_2 \\ (\rho e + p) v_1 \end{pmatrix}\,, \quad \mathbf{f}_2(\mathbf{u}) = \begin{pmatrix} \rho v_2 \\ \rho v_1 v_2 \\ \rho v_2^2 + p \\ (\rho e + p) v_2 \end{pmatrix}\,. \tag{1.21}$$

Setting for $1 \le i, k \le s$

$$[\mathbf{J}_1(\mathbf{u})]_{ik} \equiv \frac{\partial f_{1i}}{\partial u_k}(\mathbf{u}) \, , \qquad [\mathbf{J}_2(\mathbf{u})]_{ik} \equiv \frac{\partial f_{2i}}{\partial u_k}(\mathbf{u}) \, , \tag{1.22}$$

it can be easily shown that

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{\gamma-3}{2} v_1^2 + \frac{\kappa}{2} v_2^2 & (3-\gamma) v_1 & -\kappa v_2 & \kappa \\ -v_1 v_2 & v_2 & v_1 & 0 \\ -\gamma e v_1 + \kappa v_1 |\mathbf{v}|^2 & \gamma e - \frac{\kappa}{2}(3v_1^2 + v_2^2) & -\kappa v_1 v_2 & \gamma v_1 \end{pmatrix}$$

and

$$\mathbf{J}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -v_1 v_2 & v_2 & v_1 & 0 \\ \frac{\gamma-3}{2} v_2^2 + \frac{\kappa}{2} v_1^2 & -\kappa v_1 & (3-\gamma) v_2 & \kappa \\ -\gamma e v_2 + \kappa v_2 |\mathbf{v}|^2 & -\kappa v_1 v_2 & \gamma e - \frac{\kappa}{2}(3v_2^2 + v_1^2) & \gamma v_2 \end{pmatrix}$$

where $\kappa \equiv \gamma - 1$. Now, for any $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ it can be shown (Fernandez [1988]) that the matrix

$$\mathbf{P}(\mathbf{u}, \boldsymbol{\alpha}) \equiv \alpha_1 \mathbf{J}_1(\mathbf{u}) + \alpha_2 \mathbf{J}_2(\mathbf{u}) \tag{1.23}$$

has for all physically relevant values (i.e. for $\rho > 0$) four real eigenvalues, namely,

$$\begin{aligned} \lambda_1 = \lambda_2 &= \alpha_1 v_1 + \alpha_2 v_2 \, , \\ \lambda_3 &= \lambda_1 + a\sqrt{\alpha_1^2 + \alpha_2^2} \, , \\ \lambda_4 &= \lambda_1 - a\sqrt{\alpha_1^2 + \alpha_2^2} \, , \end{aligned} \tag{1.24}$$

where a is the local sound speed in isentropic processes defined by

$$a^2 \equiv \gamma \frac{p}{\rho} \, . \tag{1.25}$$

Then $\mathbf{P}$ can be shown to be a diagonalizable matrix:

$$\mathbf{P} = \mathbf{C} \, \boldsymbol{\Lambda} \, \mathbf{C}^{-1} \, , \tag{1.26}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ and (see WADA ET AL. [1988], FERNANDEZ [1988])

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & \dfrac{1}{2a^2} & \dfrac{1}{2a^2} \\ v_1 & \eta_2 & \dfrac{v_1 + a\eta_1}{2a^2} & \dfrac{v_1 - a\eta_1}{2a^2} \\ v_2 & -\eta_1 & \dfrac{v_2 + a\eta_2}{2a^2} & \dfrac{v_2 - a\eta_2}{2a^2} \\ \dfrac{|\mathbf{v}|^2}{2} & \eta_2 v_1 - \eta_1 v_2 & \dfrac{H + a\,\boldsymbol{\eta}\cdot\mathbf{v}}{2a^2} & \dfrac{H - a\,\boldsymbol{\eta}\cdot\mathbf{v}}{2a^2} \end{pmatrix}$$

while

$$\mathbf{C}^{-1} = \begin{pmatrix} 1 - \dfrac{\kappa|\mathbf{v}|^2}{2\,a^2} & \dfrac{\kappa}{a^2}v_1 & \dfrac{\kappa}{a^2}v_2 & -\dfrac{\kappa}{a^2} \\ \eta_1 v_2 - \eta_2 v_1 & \eta_2 & -\eta_1 & 0 \\ -a\,\boldsymbol{\eta}\cdot\mathbf{v} + \dfrac{\kappa|\mathbf{v}|^2}{2} & a\eta_1 - \kappa v_1 & a\eta_2 - \kappa v_2 & \kappa \\ a\,\boldsymbol{\eta}\cdot\mathbf{v} + \dfrac{\kappa|\mathbf{v}|^2}{2} & -a\eta_1 - \kappa v_1 & -a\eta_2 - \kappa v_2 & \kappa \end{pmatrix}.$$

Here,

$$H \equiv \frac{|\mathbf{v}|^2}{2} + \frac{a}{\kappa} \tag{1.27}$$

is the so-called **enthalpy** and

$$\eta_1 \equiv \frac{\alpha_1}{\sqrt{\alpha_1^2 + \alpha_2^2}}\,, \qquad \eta_2 \equiv \frac{\alpha_2}{\sqrt{\alpha_1^2 + \alpha_2^2}}\,, \qquad \kappa \equiv \gamma - 1\,. \tag{1.28}$$

Hence the system of Euler equations is a hyperbolic one (see also Section 2.1). The same result holds also for $d = 3$ (see WADA ET AL. [1988]).

1.1.2 The p-system

Consider the easiest model of one-dimensional gas dynamics for isentropic processes in Lagrangian coordinates[†] in the form of the

[†] Due to different coordinates, the operators $\frac{\partial}{\partial t}$, $\frac{\partial}{\partial x}$ do not coincide with the operators denoted by the same symbols in Euler coordinates.

so-called **p-system** (see for example LANDAU AND LIFSHITZ [1959] and KRÖNER [1996]):

$$\begin{aligned} \frac{\partial \sigma}{\partial t} - \frac{\partial v}{\partial x} &= 0\,, \\ \frac{\partial v}{\partial t} + \frac{\partial p(\sigma)}{\partial x} &= 0\,. \end{aligned} \tag{1.29}$$

Here, $\sigma = \sigma(t,x)$, $\sigma \equiv 1/\rho > 0$, is the specific volume, $v = v(t,x)$ is the velocity and $p = p(\sigma)$ is the pressure.

Denoting

$$\mathbf{w} = \begin{pmatrix} \sigma \\ v \end{pmatrix}\,, \qquad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} -v \\ p(\sigma) \end{pmatrix}\,, \tag{1.30}$$

one can write (1.29) as a system of two equations

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{w})}{\partial x} = \mathbf{0}\,. \tag{1.31}$$

The system (1.31) is hyperbolic provided that $p'(\sigma) < 0$. Indeed, the Jacobian matrix $\mathbf{J} = \left(\frac{\partial f_i}{\partial w_j}\right)$ is of the form

$$\mathbf{J} = \begin{pmatrix} 0 & -1 \\ p'(\sigma) & 0 \end{pmatrix}$$

and therefore, under the assumption $p'(\sigma) < 0$,[‡] $\mathbf{J}$ has two distinct real eigenvalues:

$$\lambda_1 = \sqrt{-p'(\sigma')}\,, \qquad \lambda_2 = -\sqrt{-p'(\sigma)}\,.$$

Hence, the system (1.31) is (strictly) hyperbolic.

Note that under the setting (justified by $(1.29)_1$)

$$v \equiv \frac{\partial u}{\partial t}\,, \quad \sigma \equiv \frac{\partial u}{\partial x}\,, \quad p(\sigma) \equiv -a(\sigma)\,,$$

the p-system (1.29) can be considered as a non-linear wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} a\left(\frac{\partial u}{\partial x}\right) = 0\,, \tag{1.32}$$

with $a'(\sigma) > 0$. In the next example we will show that also in d space dimensions there is a connection between scalar hyperbolic equations of second order of the type analogous to (1.32) and some hyperbolic systems.

‡ For example, $p(\sigma) = c\sigma^{-\gamma}$ for perfect isentropic gas ($c > 0$, $\gamma > 1$).

1.1.3 Scalar hyperbolic equation of second order

Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded domain, $\partial\Omega \in C^{0,1}$. Let $T > 0$, $I \equiv (0,T)$ and $Q_T \equiv I \times \Omega$. We consider the scalar hyperbolic equation of second order

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x_i}\big(a_i(\nabla u)\big) = 0 \qquad \text{in } Q_T\,, \tag{1.33}$$

supposing that there exists a function $\vartheta \in C^2(\mathbb{R}^d)$, called **potential** to $\boldsymbol{a} \equiv (a_1, \ldots, a_d)$, such that

$$\frac{\partial \vartheta}{\partial \xi_i}(\boldsymbol{\xi}) = a_i(\boldsymbol{\xi}) \qquad \forall\, \boldsymbol{\xi} \in \mathbb{R}^d\,, \tag{1.34}$$

for $i = 1, \ldots, d$, and

$$\frac{\partial^2 \vartheta(\boldsymbol{\xi})}{\partial \xi_i \partial \xi_j}\eta_i\eta_j \geq \alpha|\boldsymbol{\eta}|^2 \qquad \forall\, \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d \tag{1.35}$$

for a certain positive constant $\alpha \in \mathbb{R}$.

Chapter 4 is devoted to the detailed study of an initial-boundary value problem for the equation (1.33) (with a right-hand side f). Here we just show that under the assumptions (1.34)–(1.35) the equation (1.33) can be considered as a hyperbolic system of $(d+1)$ equations.

To this end, let us put $v_0 \equiv \frac{\partial u}{\partial t}$ and $v_j \equiv \frac{\partial u}{\partial x_j}$. We will also use the notation $\widehat{\mathbf{v}} \equiv (v_1, \ldots, v_d)$ and $\mathbf{v} \equiv (v_0, \widehat{\mathbf{v}})$. Then, (1.33) can be written as

$$\begin{aligned} \frac{\partial v_0}{\partial t} - \frac{\partial}{\partial x_j} a_j(\widehat{\mathbf{v}}) &= 0\,, \\ \frac{\partial v_j}{\partial t} - \frac{\partial v_0}{\partial x_j} &= 0\,, \quad 1 \leq j \leq d\,. \end{aligned} \tag{1.36}$$

Applying the chain rule to $\frac{\partial a_j}{\partial x_j}$, one can write the system (1.36) [of $(d+1)$ equations] in a matrix form:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{J}_k \frac{\partial \mathbf{v}}{\partial x_k} = 0\,, \tag{1.37}$$

where

$$\mathbf{J}_k \equiv \begin{pmatrix} 0 & -\frac{\partial a_k}{\partial \xi_1} & -\frac{\partial a_k}{\partial \xi_2} & \dots & -\frac{\partial a_k}{\partial \xi_d} \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow (k+1)^{\mathrm{st}}, \\ \\ \\ \end{array} \tag{1.38}$$

$k = 1, \dots, d$. Recall that the hyperbolicity of the system (1.37) will be proved as soon as we show that for any constant $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d) \in \mathbb{R}^d$ all the eigenvalues of the matrix

$$\mathbf{J} \equiv \eta_k \mathbf{J}_k \tag{1.39}$$

are real and the matrix itself is diagonalizable.

Denoting $S_i \equiv \eta_k \frac{\partial a_k}{\partial \xi_i}$ we see that

$$\mathbf{J} = \begin{pmatrix} 0 & -S_1 & -S_2 & \dots & -S_d \\ -\eta_1 & 0 & 0 & \dots & 0 \\ -\eta_2 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\eta_d & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{1.40}$$

Now it is not difficult to see (e.g. by induction) that

$$\det(\lambda \mathbf{I} - \mathbf{J}) = \lambda^{d+1} - \lambda^{d-1}\big(\eta_i S_i\big) \tag{1.41}$$

and therefore the eigenvalues λ_i of $\mathbf{J}$ are of the form

$$\begin{aligned} \lambda_1 = \dots = \lambda_{d-1} = 0\,, \\ \lambda_d^2 = \lambda_{d+1}^2 = \eta_i S_i\,. \end{aligned} \tag{1.42}$$

Lemma 1.43 *Under the assumptions (1.34)–(1.35) all the eigenvalues λ_i are real and the matrix $\mathbf{J}$ is diagonalizable, i.e., the system (1.37) is hyperbolic.*

PROOF : If $\eta_i = 0$ for all $i = 1, \dots, d$, then $\lambda_1 = \dots = \lambda_{d+1} = 0$ and $\mathbf{J} = \mathbf{0}$, i.e. $\mathbf{J}$ is diagonal.

Therefore we assume without loss of generality that $|\boldsymbol{\eta}| > 0$. Then, using the definition of S_i, (1.34) and (1.35), we get

$$\eta_i S_i = \eta_i \eta_k \frac{\partial^2 \vartheta}{\partial \xi_i \partial \xi_k} \geq \alpha |\boldsymbol{\eta}|^2 > 0\,. \tag{1.44}$$

Hence, all eigenvalues λ_i are real, $\lambda_d = \sqrt{\eta_i S_i} \neq -\sqrt{\eta_i S_i} = \lambda_{d+1}$.

To prove that $\mathbf{J}$ is diagonalizable we recall that $\mathbf{J}$ is diagonalizable if and only if $\dim N_\lambda = m(\lambda)$, where $N_\lambda \equiv \{\mathbf{x}; \mathbf{J}\mathbf{x} = \lambda\mathbf{x}\}$ and $m(\lambda)$ is an algebraic multiplicity of λ as a root of the characteristic polynomial $\det(\lambda\mathbf{I} - \mathbf{J})$ (cf. (1.41)). Since trivially $\dim N_\lambda \le m(\lambda)$ and $\dim N_\lambda \ge 1$ for any eigenvalue λ, it remains to prove $\dim N_0 = d - 1$.

Now, if $\mathbf{x} = (x_0, \dots, x_d)$ then from the form of the matrix $\mathbf{J}$ it is clear that $\mathbf{x} \in N_0$ if and only if $x_0 = 0$ and $\sum_{i=1}^d x_i S_i = 0$. Since there is j_0 such that $S_{j_0} \ne 0$ (otherwise we get a contradiction with (1.44)), the latter equation reduces to

$$x_{j_0} = -\frac{1}{S_{j_0}} \sum_{i \ne j_0} x_i S_i\,,$$

giving us the possibility to choose arbitrarily $(d-1)$ coordinates of $\mathbf{x}$ on the right-hand side of this expression. Thus, $\dim N_0 = d - 1$ and the proof is complete. ■

1.1.4 Compressible non-Newtonian liquids undergoing isothermal processes

In Section 1.1.1 the considered stress tensor $\mathbf{T}$ was determined by pressure, i.e.,

$$\mathbf{T} = -p(\rho, \theta)\,\mathbf{I}\,,$$

(cf. $(1.8)_1$, (1.13)). If the viscous effects taken into consideration are substantial, the dependence of $\mathbf{T}$ on other quantities, say $\nabla\mathbf{v}$, $\nabla\theta$, is supposed. In this case,[§]

$$\mathbf{T} = \widehat{\mathbf{T}}(\rho, \theta, \nabla\theta, \nabla\mathbf{v})\,.$$

However, in case the motion of the material is **isothermal**, i.e. the temperature $\theta = \theta_0 > 0$ is constant, the tensor function $\widehat{\mathbf{T}}$ does not depend on θ and $\nabla\theta$. Thus,

$$\mathbf{T} = \overline{\mathbf{T}}(\rho, \nabla\mathbf{v})\,. \tag{1.45}$$

Consequently, the equations (1.1), (1.2) are not coupled with (1.3) and can be considered separately. In other words: once having ρ,

[§] For simplicity, we suppose that $\mathbf{T}$ at each point (t, x) depends only on the values of ρ, θ, $\nabla\theta$ and $\nabla\mathbf{v}$ at the same point (t, x). Consequently, any impact of history or non-local effects are not allowed.

$\mathbf{v}$ determined from (1.1), (1.2), one can use (1.3) to calculate the remaining thermodynamical quantities.

Now, taking into account the **principle of material frame indifference** (cf. for example TRUESDELL [1991]), one can show that (1.45) reduces to

$$\mathbf{T} = -p(\rho)\,\mathbf{I} + \widetilde{\mathbf{T}}(\rho, \mathbf{e}) \,, \tag{1.46}$$

where $2\,\mathbf{e} = 2\,\mathbf{e}(\mathbf{v}) \equiv \nabla\mathbf{v} + (\nabla\mathbf{v})^T$ is the symmetric part of the velocity gradient $\nabla\mathbf{v}$.

In this book we will study a special form of (1.46), namely

$$\mathbf{T} = -p(\rho)\,\mathbf{I} + \boldsymbol{\tau}^E \,, \tag{1.47}$$

with $\boldsymbol{\tau}^E$ given by

$$\boldsymbol{\tau}^E = \boldsymbol{\tau}(\mathbf{e}) \,. \tag{1.48}$$

Here, $\boldsymbol{\tau} : \mathbb{R}^{d^2}_{\text{sym}} \longrightarrow \mathbb{R}^{d^2}_{\text{sym}}$ is a given continuous function and $\mathbb{R}^{d^2}_{\text{sym}} \equiv \big\{\mathbf{M} \in \mathbb{R}^d \times \mathbb{R}^d; M_{ij} = M_{ji},\, i,j = 1,\dots,d\big\}$.

Assuming that a liquid (gas) obeys the state equation (1.17), we obtain that (in the isothermal case) the pressure is a linear function of ρ, i.e.,

$$p(\rho) = \beta\rho \,, \qquad \beta = R\theta_0 > 0 \,. \tag{1.49}$$

Under the assumptions (1.47)–(1.49), the system (1.1), (1.2) reads

$$\frac{\partial\rho}{\partial t} + \frac{\partial(\rho v_j)}{\partial x_j} = 0 \,, \tag{1.50}$$

$$\frac{\partial(\rho v_i)}{\partial t} + \frac{\partial(\rho v_i v_j)}{\partial x_j} = -\beta\frac{\partial\rho}{\partial x_i} + \frac{\partial\tau_{ij}(\mathbf{e})}{\partial x_j} + \rho f_i \,, \tag{1.51}$$

for $i = 1,\dots,d$. The left-hand side of (1.51) is equal to $\rho\frac{\partial v_i}{\partial t} + \rho v_j \frac{\partial v_i}{\partial x_j}$ due to (1.50).

In order to develop a mathematical theory (see Section 5.5), we will assume that there exist constants C_1, $C_2 > 0$ and parameters $p > 1$ and $q \in [p-1, p)$ such that for all $\boldsymbol{\eta} \in \mathbb{R}^{d^2}_{\text{sym}}$ the p**-coercivity** condition

$$\boldsymbol{\tau}(\boldsymbol{\eta}) \cdot \boldsymbol{\eta} \geq C_1 |\boldsymbol{\eta}|^p \,, \tag{1.52}$$

and the q**-growth** condition

$$|\boldsymbol{\tau}(\boldsymbol{\eta})| \leq C_2\big(1 + |\boldsymbol{\eta}|\big)^q \tag{1.53}$$

hold. As usual, $|\boldsymbol{\eta}| = (\eta_{ij}\eta_{ij})^{1/2}$ and $\boldsymbol{\tau} \cdot \boldsymbol{\eta} = \tau_{ij}\eta_{ij}$ for $\boldsymbol{\tau}$, $\boldsymbol{\eta} \in \mathbb{R}^{d^2}_{\text{sym}}$.

1.1.5 *Incompressible non-Newtonian fluids undergoing isothermal processes*

If a material is incompressible, i.e. $\rho(t,x) = \rho_0 > 0$ for all $(t,x) \in [0,T) \times \Omega$, and undergoes an isothermal process, then

- we get from (1.45) that

$$\mathbf{T} = -\pi\,\mathbf{I} + \boldsymbol{\tau}^E\,, \tag{1.54}$$

  where π is the so-called **undetermined pressure** and $\boldsymbol{\tau}^E$ is the **extra stress** tensor;
- the system (1.50)–(1.51) reads

$$\operatorname{div}\mathbf{v} = \frac{\partial v_i}{\partial x_i} = 0\,, \tag{1.55}$$

$$\rho_0 \frac{\partial v_i}{\partial t} + \rho_0 v_j \frac{\partial v_i}{\partial x_j} = -\frac{\partial \pi}{\partial x_i} + \frac{\partial \tau^E_{ij}}{\partial x_j} + \rho_0 f_i\,, \tag{1.56}$$

  for $i = 1, \dots, d$.

We will assume that the extra stress $\boldsymbol{\tau}^E$ is given by the sum of two symmetric tensor functions of $\mathbf{e} = \mathbf{e}(\mathbf{v})$. This means that we have for all $(t,x) \in [0,T) \times \Omega$,

$$\boldsymbol{\tau}^E(t,x) = \boldsymbol{\tau}\big(\mathbf{e}(\mathbf{v}(t,x))\big) + \boldsymbol{\sigma}\big(\mathbf{e}(\mathbf{v}(t,x))\big)\,, \tag{1.57}$$

or equivalently

$$\boldsymbol{\tau}^E = \boldsymbol{\tau}(\mathbf{e}) + \boldsymbol{\sigma}(\mathbf{e})\,. \tag{1.58}$$

Similarly to Section 1.1.4, the following assumptions are imposed on $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}$: for certain $p > 1$, $q \in [p-1, p)$, C_1, $C_2 > 0$, we have

$$\boldsymbol{\tau}(\boldsymbol{\eta}) \cdot \boldsymbol{\eta} \geq C_1 |\boldsymbol{\eta}|^p\,, \tag{1.59}$$

$$\boldsymbol{\sigma}(\boldsymbol{\eta}) \cdot \boldsymbol{\eta} \geq 0\,, \tag{1.60}$$

$$|\boldsymbol{\tau}(\boldsymbol{\eta}) + \boldsymbol{\sigma}(\boldsymbol{\eta})| \leq C_2\big(1 + |\boldsymbol{\eta}|\big)^q \tag{1.61}$$

for all $\boldsymbol{\eta} \in \mathbb{R}^{d^2}_{\mathrm{sym}}$.

Notice that the tensor $\boldsymbol{\tau} + \boldsymbol{\sigma}$ satisfies (1.52)–(1.53) and it might not be clear why the decomposition (1.58) is introduced. Actually, we will use (1.58) only in such parts of the mathematical theory where the assumptions on $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}$ will be strengthened. Roughly speaking, we will suppose that $\boldsymbol{\tau}$ is well behaved and $\boldsymbol{\sigma}$ is controlled by $\boldsymbol{\tau}$. By 'well behaved' we mean the existence of a strictly convex potential $U : \mathbb{R}^{d^2}_{\text{sym}} \longrightarrow \mathbb{R}$ of $\boldsymbol{\tau}$, such that

$$\frac{\partial U(\mathbf{e})}{\partial e_{ij}} = \tau_{ij}(\mathbf{e}) , \qquad i,j = 1,\dots,d .$$

Appropriate assumptions put on U will in particular imply (1.59). For more details see Section 5.1.

Some interesting examples of $\boldsymbol{\tau}^E$ of the form (1.58) can be found in what follows.

Example 1.62 (Stokes' law) If the dependence of $\boldsymbol{\tau}$ on $\mathbf{e}$ is linear, i.e.,

$$\boldsymbol{\tau}(\mathbf{e}) = 2\,\nu\,\mathbf{e} , \quad \nu > 0 , \tag{1.63}$$

and $\boldsymbol{\sigma} \equiv \mathbf{0}$, then the system (1.55)–(1.56) turns into the well-known **Navier-Stokes system**

$$\operatorname{div} \mathbf{v} = 0 , \tag{1.64}$$

$$\rho_0 \frac{\partial v_i}{\partial t} + \rho_0 v_j \frac{\partial v_i}{\partial x_j} = -\frac{\partial \pi}{\partial x_i} + \nu \Delta v_i + \rho_0 f_i , \tag{1.65}$$

for $i = 1, 2, \dots, d$. Let us notice that in this case

$$\boldsymbol{\tau}(\boldsymbol{\eta}) \cdot \boldsymbol{\eta} = 2\,\nu |\boldsymbol{\eta}|^2 .$$

Therefore, the condition (1.59) is satisfied with $p = 2$.

Definition 1.66 *An incompressible fluid, the behaviour of which is characterized by Stokes' law (1.63), is called* **Newtonian fluid**. *Fluids that cannot be adequately described by (1.63) are usually called* **non-Newtonian fluids**.

Due to the negative character of the definition of non-Newtonian fluids, it would be useful:

- to characterize main points of deviance from Newtonian behaviour,
- to classify the models of some non-Newtonian fluids.

This book is not intended to be an introduction to the mechanics of non-Newtonian fluids. For this purpose, we refer the reader to SCHOWALTER [1978], HUILGOL [1975] and RAJAGOPAL [1993]. Nevertheless, we would like at least to specify which of the basic properties of non-Newtonian behaviour can be captured by the model characterized by (1.58).

Let us recall (see e.g. RAJAGOPAL [1993]) that the main points of non-Newtonian behaviour are:

1. the ability of the fluid to shear thin or shear thicken in shear flows;
2. the presence of non-zero normal stress differences in shear flows;
3. the ability of the fluid to yield stress;
4. the ability of the fluid to exhibit stress relaxation;
5. the ability of the fluid to creep.

A non-Newtonian fluid can possess just one or all of the above listed characteristics.

The model of fluid described by (1.58) exhibits mainly the first property, while it cannot predict 'elastic' phenomena 2–4. This means that for a fluid given by (1.58) the dominant departure from the Newtonian behaviour is shear thinning or shear thickening, while the other features are not captured. For this reason we will discuss just the first phenomenon in more detail.

Consider a steady shear flow for which

$$\mathbf{v} = \big(v_1(x_2), 0, 0\big) .$$

Setting $\kappa \equiv |\frac{d}{dx_2} v_1(x_2)|$ we can define the so-called **generalized** (or **apparent**) **viscosity** μ_g by

$$\mu_g(\kappa) \equiv \frac{\tau_{12}(\kappa) + \sigma_{12}(\kappa)}{\kappa} . \tag{1.67}$$

Definition 1.68 *If the generalized viscosity μ_g defined in (1.67) is an increasing function of κ, then the fluid described by (1.58) is called* **shear thickening fluid**. *If μ_g is a decreasing function of κ, the fluid is called* **shear thinning fluid**.

Note that for Stokes' law (1.63) we have $\mu_g(\kappa) = \nu$ for all κ.

Example 1.69 (Generalized Newtonian fluids and power-law fluids) Let $\boldsymbol{\tau}$ be given by

$$\boldsymbol{\tau}(\mathbf{e}) = 2\mu\big(|\mathbf{e}|^2\big)\, \mathbf{e} = 2\widetilde{\mu}(\mathbf{e})\, \mathbf{e} \tag{1.70}$$

and $\boldsymbol{\sigma} \equiv 0$. Then, the potential U is defined by

$$U(\mathbf{e}) = \int_0^{|\mathbf{e}|^2} \mu(s)\, ds . \tag{1.71}$$

If, in particular, we take

$$\mu(s) = \nu_0 s^{\frac{r}{2}} , \quad \nu_0 > 0 , \tag{1.72}$$

then

$$\begin{aligned} U(\mathbf{e}) &= \frac{2\nu_0}{r+2}|\mathbf{e}|^{r+2}\,, \\ \boldsymbol{\tau}(\boldsymbol{\eta})\cdot\boldsymbol{\eta} &= 2\,\nu_0|\boldsymbol{\eta}|^r\boldsymbol{\eta}\cdot\boldsymbol{\eta} = 2\,\nu_0|\boldsymbol{\eta}|^{r+2}\,, \end{aligned}$$

and we see from (1.59) that $p = r+2$. Therefore $p \in (1,+\infty)$ if and only if $r \in (-1,+\infty)$. Further, for the shear flow $\mu_g(\kappa) = \mu(\kappa)$, $\frac{d\mu}{d\kappa} < 0$ if $r \in (-1,0)$ and $\frac{d\mu}{d\kappa} > 0$ if $r > 0$. In other words, the model (1.70) with (1.72) captures the shear thinning fluid if $r \in (-1,0)$ (or $p \in (1,2)$), and captures the shear thickening fluid if $r > 0$ (or $p > 2$). The case $r = 0$ (or $p = 2$) corresponds to the Newtonian fluid.

The fluids characterized by (1.70) are called **generalized Newtonian fluids** (even if they are non-Newtonian ones). The fluids described by (1.70) and (1.72) are called **power-law fluids**.

We refer to Málek, Rajagopal and Růžička [1995] for an exhaustive, but not complete list of literature, where the power-law fluids (1.70), (1.72) are used in several fields of chemistry, glaciology, biology, geology, etc.

Example 1.73 (Various variants of power-law fluids) Despite its simple structure, the model (1.70) includes submodels with a great deal of disparity. To illustrate this fact, we will investigate apparently almost identical types of $\boldsymbol{\tau}$ in a shear flow. Let us consider

$$\begin{aligned} &\text{(a)} && \boldsymbol{\tau}^{(1)}(\mathbf{e}) = 2\,\nu_0|\mathbf{e}|^r\mathbf{e}\,, \\ &\text{(b)} && \boldsymbol{\tau}^{(2)}(\mathbf{e}) = 2\,\nu_0\big(1+|\mathbf{e}|\big)^r\mathbf{e}\,, \\ &\text{(c)} && \boldsymbol{\tau}^{(3)}(\mathbf{e}) = 2\,\nu_0\big(1+|\mathbf{e}|^2\big)^{r/2}\mathbf{e}\,, \\ &\text{(d)} && \boldsymbol{\tau}^{(3+i)}(\mathbf{e}) = 2\,\nu_\infty\mathbf{e} + \boldsymbol{\tau}^{(i)}(\mathbf{e})\,, \quad i = 1,2,3\,, \end{aligned} \tag{1.74}$$

where ν_0 and ν_∞ are positive constants related to the limits of $\mu_g(\kappa)$ when $\kappa \to 0$ and $\kappa \to \infty$, respectively.

Let first $r \in (-1,0)$. Figure 1.1 depicts the graphs of μ_g corresponding to the tensors $\boldsymbol{\tau}^{(i)}$, $i = 1,\dots,6$, defined in (1.74).

It is worth observing that the value of parameter p in (1.59) varies for different models in (1.74); while $p = r+2$ for the cases (a), (b), (c), we see that $p = 2$ for all cases in (d).

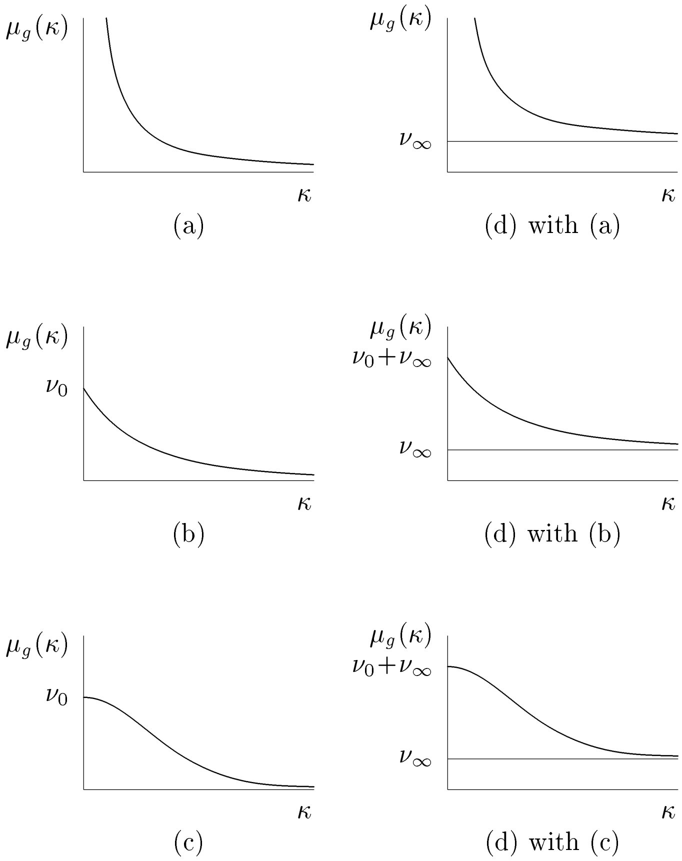

Figure 1.1 *The graphs of* μ_g *for* $\boldsymbol{\tau}^{(i)}$, $i = 1, \ldots, 6$, *with* $r \in (-1, 0)$.

We will see in Chapter 5 that some results cannot be proved for all values of $p > 1$. More often, we will get some restriction p_0 on the value of p from below such that the corresponding result is valid for all $p > p_0$. It can happen that $r + 2 < p_0 < 2$. Then the result is valid for the fluid described by (1.74)(d), while our approach fails for any of the models (a)–(c).

In fact, the model (1.74)(d) can be understood in the form of (1.58). For example, we can decompose $\boldsymbol{\tau}^{(4)}(\mathbf{e}) = \boldsymbol{\tau}(\mathbf{e}) + \boldsymbol{\sigma}(\mathbf{e})$, where

$$\boldsymbol{\tau}(\mathbf{e}) = 2\,\nu_\infty \mathbf{e}\,, \qquad \nu_\infty > 0\,, \tag{1.75}$$

and

$$\boldsymbol{\sigma}(\mathbf{e}) = 2\,\nu_0 |\mathbf{e}|^r \mathbf{e}\,, \qquad \nu_0 > 0\,. \tag{1.76}$$

It can be easily verified that in this case $r \geq -1$ can even attain the value -1. In particular, the case $r = -1$ is of special interest from the point of view of applications (see (1.82) for example).

Let us now consider the case $r > 0$. Then we can draw the pictures for the models (1.74a, b, c), see Figure 1.2.

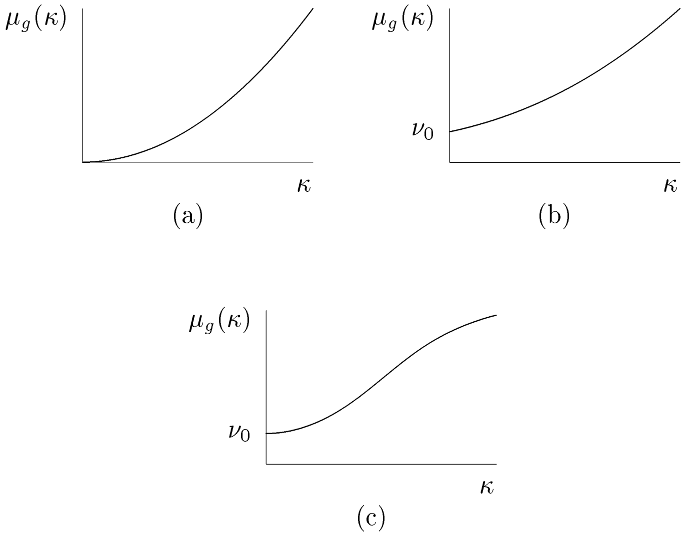

Figure 1.2 *The graphs of* μ_g *for* $\boldsymbol{\tau}^{(i)}$, $i = 1,\ 2,\ 3$, *with* $r \in (0,1)$.

We are primarily interested in fluids that have a non-zero "zero shear rate viscosity". Figure 1.2 (a) reveals that the viscosity function vanishes with the shear rate (when $\kappa \to 0$) and thus we shall not consider this possibility any further. The particular case of (d) ($r > 0$), namely

$$\boldsymbol{\tau}(\mathbf{e}) = 2\,\nu_\infty \mathbf{e} + 2\,\nu_0 |\mathbf{e}|^r \mathbf{e}\,, \qquad \nu_0, \nu_\infty > 0\,, \tag{1.77}$$

was introduced in the mathematical literature by LADYZHENSKAYA [1969] and the corresponding system (1.55), (1.56) and (1.77) is sometimes called the **modified Navier-Stokes system**.

Using (1.70) and (1.71) it is easy to observe that the potentials $U^{(i)}$ corresponding to the tensors $\boldsymbol{\tau}^{(i)}$ defined in (1.74) for $r > -1$ and $i = 1, \dots, 6$, are as follows:

$$\begin{aligned}
&\text{(a)} \qquad U^{(1)}(\mathbf{e}) = \frac{2\,\nu_0}{r+2}|\mathbf{e}|^{r+2}\,,\\
&\text{(b)} \qquad U^{(2)}(\mathbf{e}) = \frac{2\,\nu_0}{r+2}\big(1+|\mathbf{e}|\big)^{r+2} - \frac{2\,\nu_0}{r+1}\big(1+|\mathbf{e}|\big)^{r+1}\\
&\qquad\qquad\qquad + \frac{2\,\nu_0}{(r+1)(r+2)}\,,\\
&\text{(c)} \qquad U^{(3)}(\mathbf{e}) = \frac{2\,\nu_0}{r+2}\Big[\big(1+|\mathbf{e}|^2\big)^{\frac{r+2}{2}} - 1\Big]\,,\\
&\text{(d)} \qquad U^{(3+i)}(\mathbf{e}) = \nu_\infty|\mathbf{e}|^2 + U^{(i)}(\mathbf{e})\,, \quad i = 1,2,3\,.
\end{aligned}$$

In the last part of this section we will give three examples of models which are widely used in recent years as up-to-date ones in various fields of applied sciences. The first example presents the latest model for the flow of glacier; the second example has been proposed in blood rheology for modelling the flow of blood through arteries. Finally, the last model is used in geology for describing the dynamics of tectonic plates in the Earth's mantle.

In connection with the introduction of these models we wish to emphasize one more point here. As we shall see below, all three models mentioned above belong to the class described by (1.55)–(1.61) only for a special choice of values of parameters. It is presumed, however, that the mathematical theory developed in Chapter 5 could serve as a starting point of mathematical study of the models, presented in the following examples.

Example 1.78 (Glacier ice in creeping flow) Based on experimental tests performed by KJARTANSON [1986], KJARTANSON ET AL. [1988] and in agreement with experimental results of VAN DER VEEN AND WHILLANS [1990], Man together with his co-workers (see MAN AND SUN [1987] and references therein) proposed that the system (1.55)–(1.56) with the extra stress $\boldsymbol{\tau}^E$ given by

$$\boldsymbol{\tau}^E = \mu|\mathbf{A}_1|^r\mathbf{A}_1 + \alpha_1\mathbf{A}_2 + \alpha_2\mathbf{A}_1^2 \tag{1.79}$$

is a reasonable model for the creeping flow of ice. The tensors $\mathbf{A}_1$, $\mathbf{A}_2$ are the first two of the so-called **Rivlin-Ericksen tensors**, r, μ, α_1, α_2 are material constants, $\mu > 0$, $r \approx -\frac{2}{3}$. Let us recall (see for example TRUESDELL [1991]) that

$$\mathbf{A}_1 = 2\,\mathbf{e}$$

and

$$\mathbf{A}_2 = \frac{d}{dt}\mathbf{A}_1 + \mathbf{A}_1(\nabla\mathbf{v}) + (\nabla\mathbf{v})^T\mathbf{A}_1 \,.$$

Notice that on putting $\alpha_1 = \alpha_2 = 0$ in (1.79) we obtain (1.70), (1.72). On the other hand, (1.79) with general α_1, α_2 cannot be included into the class described by (1.58)–(1.61) due to the dependence of $\boldsymbol{\tau}^E$ not only on $\mathbf{A}_1$ but also on $\frac{d}{dt}\mathbf{A}_1$. In fact, the model (1.79) with $r = 0$ describes what is called the **second grade fluid**. We refer to DUNN AND RAJAGOPAL [1995] for an exposition on n-**grade** fluids and for further references.

Example 1.80 (Blood flow) Experimental tests reveal that blood exhibits non-Newtonian phenomena such as shear thinning, creep and stress relaxation. In order to include all these features in the model, YELESWARAPU ET AL. [1994] proposed the so-called **generalized Oldroyd-B model**. The constitutive equation of that model has the form

$$\begin{aligned}\boldsymbol{\tau}^E + \lambda_1 \left[\frac{d}{dt}\boldsymbol{\tau}^E - (\nabla\mathbf{v})\,\boldsymbol{\tau}^E - \boldsymbol{\tau}^E(\nabla\mathbf{v})^T\right]& \\ = \mu(\mathbf{e})\,\mathbf{e} + \lambda_2\left(\frac{d\mathbf{e}}{dt} - (\nabla\mathbf{v})\,\mathbf{e} - \mathbf{e}\,(\nabla\mathbf{v})^T\right)&\end{aligned} \tag{1.81}$$

with

$$\mu(\mathbf{e}) = \nu_\infty + (\nu_0 - \nu_\infty)\left[\frac{1+\ln(1+\lambda|\mathbf{e}|)}{1+\lambda|\mathbf{e}|}\right]. \tag{1.82}$$

Here $\lambda > 0$ is a material constant, λ_1, λ_2 are the relaxation and retardation times, respectively, and $\nu_0 > \nu_\infty > 0$ are the limits of $\mu(\mathbf{e})$ when $|\mathbf{e}| \to 0$ and $|\mathbf{e}| \to \infty$, respectively. (Compare with Figure 1.1.) The model (1.81)–(1.82) can be included into the class (1.55)–(1.61) only when $\lambda_1 = \lambda_2 = 0$.

Example 1.83 (Dynamics in the Earth's mantle) The study of flows in the Earth's mantle consists of thermal convection in a highly viscous fluid. As pointed out in MALEVSKY AND YUEN [1991], laboratory tests of the creep of mantle materials show a non-linear dependence of $\boldsymbol{\tau}^E$ on $\mathbf{e}$. For a description of dynamics in the planet's mantle the so-called **Boussinesq approximation** for the power-law fluid is used. For example, MALEVSKY AND YUEN [1991] investigate

$$\operatorname{div} \mathbf{v} = 0\,, \tag{1.84}$$

$$-\frac{\partial \tau^E_{ij}}{\partial x_j} + \frac{\partial \pi}{\partial x_i} = \mathrm{Ra}\,\theta \mathbf{e}_d\,, \quad 1 \le i \le d\,, \tag{1.85}$$

$$\frac{\partial \theta}{\partial t} + v_j \frac{\partial \theta}{\partial x_j} - \Delta\theta = 0\,, \tag{1.86}$$

where θ is the temperature, Ra is the so-called non-dimensional Rayleigh number and $\mathbf{e}_d \equiv (0, 0, \ldots, 1)$. Further,

$$\boldsymbol{\tau}^E(\mathbf{e}) = A|\mathbf{e}|^r \mathbf{e}\,, \qquad A > 0\,,\ r \approx -\frac{2}{3} \text{ or } -\frac{4}{5}\,. \tag{1.87}$$

More precisely, the Boussinesq approximation reads

$$\operatorname{div} \mathbf{v} = 0\,, \tag{1.88}$$

$$\rho \frac{\partial v_i}{\partial t} + \rho v_j \frac{\partial v_i}{\partial x_j} = \frac{\partial \tau^E_{ij}}{\partial x_j} - \frac{\partial \pi}{\partial x_i} - \rho\theta \mathbf{e}_d\,, \tag{1.89}$$

$$\rho \frac{\partial \theta}{\partial t} + \rho v_j \frac{\partial \theta}{\partial x_j} - \Delta\theta = 0\,, \tag{1.90}$$

which converts to (1.55)–(1.61) for $\theta = \text{const}$. We refer the reader to PADULA [1994] and HILLS AND ROBERTS [1991] for a derivation of Boussinesq approximation.

In (1.85), the term $\frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j}$ is neglected since the changes in velocity are substantially smaller with respect to the changes of temperature.

1.2 Function spaces

This book is devoted to the mathematical analysis of systems of equations described in Section 1.1. The fundamental problem of such analysis is to show the existence of solutions in appropriate function spaces. Thus, in the next section, we give a survey of their definitions and basic properties.

1.2.1 Basic elements of Banach spaces

Let X be a Banach space equipped with a norm $\|\cdot\|_X$. By X^* we denote the dual space to X consisting of all linear continuous functionals

$$\varphi : X \to \mathbb{R}\,.$$

Let $\varphi \in X^*$ and $x \in X$. Then $\langle \varphi, x\rangle_X$ denotes the value of φ at the point x or we say that the brackets $\langle \cdot, \cdot \rangle_X$ denote the duality between X and X^*. The natural norm on X^* is defined as

$$\|\varphi\|_{X^*} \equiv \sup_{\|x\|_X \le 1} |\langle \varphi, x\rangle_X|\,.$$

Let $\{x_n\} \subset X$ be a sequence of elements from X. Different types of convergence can be introduced. Let $x \in X$.

1. A sequence $\{x_n\}$ converges **strongly** to x,

$$x_n \to x \quad \text{in } X$$

   if and only if

$$\|x_n - x\|_X \to 0 \quad \text{as} \quad n \to \infty\,.$$

2. A sequence $\{x_n\}$ converges **weakly** to x,

$$x_n \rightharpoonup x \quad \text{in } X$$

   if and only if for all $\varphi \in X^*$

$$\langle \varphi, x_n\rangle_X \to \langle \varphi, x\rangle_X \quad \text{as} \quad n \to \infty\,.$$

3. Let Z be a predual space to X, i.e. Z is a Banach space satisfying $Z^* = X$. Then a sequence $\{x_n\}$ converges **weakly-*** to x,

$$x_n \overset{*}{\rightharpoonup} x \quad \text{in } X$$

   if and only if for all $\xi \in Z$

$$\langle x_n, \xi\rangle_Z \to \langle x, \xi\rangle_Z \quad \text{as} \quad n \to \infty\,.$$

Theorem 2.1 (Alaoglu) *Let X have a separable predual Z and assume the sequence $\{x_n\} \subset X$ is bounded in X. Then there exists a subsequence $\{x_{n_k}\} \subset \{x_n\}$ and $x \in X$ with*

$$x_{n_k} \overset{*}{\rightharpoonup} x \qquad \text{in } X\,.$$

PROOF : See, e.g., YOSIDA [1965, Appendix to Chapter 5]. ∎

As far as X is a reflexive separable space, the weak and weak-* convergences coincide. Consequently, for a reflexive separable Banach space X, $\overset{*}{\rightharpoonup}$ can be replaced by $\rightharpoonup$ in the previous theorem.

Let us recall that a Banach space is **reflexive** if and only if $J(X) = (X^*)^*$, where $J : X \to (X^*)^*$ is the canonical isomorphism defined by

$$\langle J(x), \varphi \rangle_{X^*} \equiv \langle \varphi, x \rangle_X \qquad \forall \varphi \in X^* .$$

Theorem 2.2 *Let $x_n \rightharpoonup x$ (or $x_n \overset{*}{\rightharpoonup} x$) in X. Then*

- *$\{x_n\}$ is bounded in X,*
- *$\|x\|_X \le \liminf_{n\to\infty} \|x_n\|_X$.*

PROOF : See for example YOSIDA [1965, Chapter 5]. ■

Let X, Y be two Banach spaces. The space X is (**continuously**) **imbedded** into Y,

$$X \hookrightarrow Y ,$$

if and only if

- $X \subset Y$,
- there exists $c > 0$ such that $\|x\|_Y \le c\,\|x\|_X$ for all $x \in X$.

The space X is **compactly imbedded** into Y,

$$X \hookrightarrow\hookrightarrow Y ,$$

if and only if

- $X \hookrightarrow Y$,
- the identity map $I : X \to Y$ is compact, i.e. $\overline{I(B)}$ is compact in Y for every bounded subset B of X.

Recall that if $X \hookrightarrow Y$ then $Y^* \hookrightarrow X^*$ and if $X \hookrightarrow\hookrightarrow Y$ then $Y^* \hookrightarrow\hookrightarrow X^*$.

1.2.2 Spaces of continuous functions

Many examples of Banach spaces will be introduced in the sequel. Most of them consist of functions (scalar or vector) defined on a domain Ω, i.e. on an open set in the Euclidean space $\mathbb{R}^d$, $d \in \mathbb{N}$. Hereafter, we will give definitions of function spaces only for scalar functions, but the definitions and also all corresponding properties of these spaces can be easily extended to vector functions. The following convention for the notation of spaces of vector functions is used in the whole book:

1. If $X(\Omega)$ denotes some space of scalar functions $u : \Omega \to \mathbb{R}$ then the space of vector functions $\mathbf{u} : \Omega \to \mathbb{R}^s$, $s \in \mathbb{N}$, for which

each component u_j, $j = 1, \dots, s$, belongs to $X(\Omega)$, is denoted by $X(\Omega)^s$.

2. Vector functions are printed in boldface (as $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$) in contrast to scalar functions which are printed in italic mode (as u, v, w).

In the case when a domain Ω is not the entire space $\mathbb{R}^d$, some information about the smoothness of boundary $\partial\Omega$ is useful. Let us therefore recall the definition of spaces of **smooth functions**.

Let Ω be a domain in $\mathbb{R}^d$. Denote $\partial\Omega \equiv \overline{\Omega} \setminus \Omega$ where, as usual, $\overline{\mathbb{R}^d} = \mathbb{R}^d$. The space of all real continuous functions $u : \overline{\Omega} \to \mathbb{R}$ for which

$$\|u\|_\infty = \|u\|_{C(\overline{\Omega})} \equiv \sup_{x \in \overline{\Omega}} |u(x)| ,$$

is finite, is denoted by $C(\overline{\Omega})$. Equipped with the norm $\|\cdot\|_\infty$, $C(\overline{\Omega})$ is a Banach space.

Let $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{N} \cup \{0\}$, be a **multi-index** and put $|\alpha| = \sum_{i=1}^d \alpha_i$. For a function $u : \Omega \to \mathbb{R}$, the symbol $D^\alpha u$ denotes partial derivatives of the order $|\alpha|$,

$$D^\alpha u \equiv \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} .$$

For $k \in \mathbb{N}$ we denote by $C^k(\overline{\Omega})$ the space of all functions u which together with their derivatives $D^\alpha u$, $|\alpha| \le k$, belong to $C(\overline{\Omega})$. The norm in $C^k(\overline{\Omega})$ is defined by

$$\|u\|_{C^k(\overline{\Omega})} \equiv \sum_{|\alpha| \le k} \|D^\alpha u\|_{C(\overline{\Omega})} .$$

Let us put

$$C^\infty(\overline{\Omega}) \equiv \bigcap_{k=0}^{\infty} C^k(\overline{\Omega}) .$$

The subspaces of $C(\overline{\Omega})$, $C^k(\overline{\Omega})$ of functions having compact support in Ω will be denoted by $C_C(\Omega)$, $C_C^k(\Omega)$, respectively, except for the space of functions from $C^\infty(\overline{\Omega})$ with compact support in Ω for which we use the notation $\mathcal{D}(\Omega)$.

Let us denote all functions from $C^\infty(\mathbb{R}^d)$ which are periodic in all directions with some period $L > 0$ by $C^\infty_{\mathrm{per}}(\Omega)$, where $\Omega = (0, L)^d$.

The space of **distributions** on Ω denoted by $\mathcal{D}'(\Omega)$, consists of all continuous linear functionals on $\mathcal{D}(\Omega)$. If $G \in \mathcal{D}'(\Omega)$ then the

distributional derivative $D^\alpha G \in \mathcal{D}'(\Omega)$ (or the derivative in the sense of distributions) of G is understood as

$$\langle D^\alpha G, \varphi \rangle_{\mathcal{D}(\Omega)} = (-1)^{|\alpha|} \langle G, D^\alpha \varphi \rangle_{\mathcal{D}(\Omega)} \qquad \forall\, \varphi \in \mathcal{D}(\Omega)\,.$$

Further, let $0 < \beta \le 1$. For $u : \overline{\Omega} \to \mathbb{R}$, put

$$[u]_\beta \equiv \sup_{\substack{x,y \in \overline{\Omega} \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|^\beta}$$

and

$$\|u\|_{C^{k,\beta}(\overline{\Omega})} \equiv \|u\|_{C^k(\overline{\Omega})} + \sum_{|\alpha| = k} [D^\alpha u]_\beta\,.$$

Then the space of **Hölder continuous functions** $C^{k,\beta}(\overline{\Omega})$ is defined as

$$C^{k,\beta}(\overline{\Omega}) \equiv \{u \in C^k(\overline{\Omega}); \|u\|_{C^{k,\beta}(\overline{\Omega})} < \infty\}\,.$$

Now we can give the 'definition' of smooth domains. Let $\Omega \subset \mathbb{R}^d$ be an open set. Roughly speaking, the boundary $\partial\Omega$ is of class $C^{k,\beta}$, $k + \beta \ge 1$,

$$\partial\Omega \in C^{k,\beta}\,,$$

if and only if '$\partial\Omega$ is $(d-1)$-dimensional $C^{k,\beta}$-manifold with Ω lying locally on one side of $\partial\Omega$'. Precise definition can be found in Nečas [1967, Chapter 1]. In Wloka [1987], the domains possessing weaker regularity properties, as segment or cone property, are introduced and discussed.

Theorem 2.3 (Imbeddings) *Let $\partial\Omega \in C^{0,1}$.*

- *Let $m \ge k \ge 0$ and $m + \beta_1 > k + \beta_2$ with $0 < \beta_i \le 1$, $i = 1, 2$. Then*
$$C^{m,\beta_1}(\overline{\Omega}) \hookrightarrow\hookrightarrow C^{k,\beta_2}(\overline{\Omega})\,. \tag{2.4}$$

- *Let $m \ge 1$. Then*
$$C^m(\overline{\Omega}) \hookrightarrow C^{m-1,1}(\overline{\Omega})\,. \tag{2.5}$$

Proof : See Kufner, John and Fučík [1977, Section 1.5]. ■

1.2.3 Lebesgue spaces

Let Ω be a domain in $\mathbb{R}^d$, $d \geq 1$, $1 \leq p \leq \infty$. We denote by $L^p(\Omega)$ the space of all measurable functions $u : \Omega \to \mathbb{R}$ for which the norm

$$\|u\|_p \equiv \left\{ \int_\Omega |u(x)|^p \, dx \right\}^{1/p}, \quad (1 \leq p < \infty)$$

or

$$\|u\|_\infty \equiv \operatorname*{ess\,sup}_{x \in \Omega} |u(x)|, \quad (p = \infty),$$

respectively, is finite. We recall that

$$\operatorname*{ess\,sup}_{x \in \Omega} |u(x)| \equiv \inf_{\substack{N \subset \Omega \\ |N|=0}} \sup_{x \in \Omega \setminus N} |u(x)|,$$

where $|N|$ denotes the Lebesgue measure of the set N.

Lemma 2.6 (Hölder's inequality) *Assume* $1/p + 1/q = 1$, $1 < p, q < \infty$ *or* $p = 1$, $q = \infty$. *Then for* $u \in L^p(\Omega)$ *and* $v \in L^q(\Omega)$

- $u\, v \in L^1(\Omega)$,
- $\|u\, v\|_1 \leq \|u\|_p \, \|v\|_q$.

PROOF : We omit trivial cases $p = 1$, $q = \infty$ and $\|u\|_p = 0$ or $\|v\|_q = 0$. Then Hölder's inequality follows from **Young's inequality**: for all a, $b \geq 0$ we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{2.7}$$

Indeed, setting $a = \frac{|u(x)|}{\|u\|_p}$ and $b = \frac{|v(x)|}{\|v\|_q}$, we get

$$\begin{aligned}\frac{\|u\, v\|_1}{\|u\|_p \, \|v\|_q} &= \int_\Omega \frac{|u(x)|}{\|u\|_p} \frac{|v(x)|}{\|v\|_q} \, dx \\ &\overset{(2.7)}{\leq} \frac{1}{p} \frac{\int_\Omega |u(x)|^p \, dx}{\|u\|_p^p} + \frac{1}{q} \frac{\int_\Omega |v(x)|^q \, dx}{\|v\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1,\end{aligned}$$

which gives the assertion. ∎

Let us note that an elementary proof of Young's inequality follows from the concavity of the logarithmic function: if a, $b > 0$ then

$$\begin{aligned}\log(ab) = \log a + \log b &= \tfrac{1}{p} \log a^p + \tfrac{1}{q} \log b^q \\ &\leq \log\left(\tfrac{a^p}{p} + \tfrac{b^q}{q}\right).\end{aligned}$$

We will frequently use an ε-version of (2.7), i.e.,

$$ab \leq \varepsilon a^p + \frac{b^q}{q(\varepsilon p)^{q/p}} = \varepsilon a^p + C(\varepsilon) b^q \,. \tag{2.8}$$

Corollary 2.9 *Let Ω be a bounded domain in $\mathbb{R}^d$ and $\infty \geq p \geq q \geq 1$. Then*

- $\|u\|_q \leq |\Omega|^{1/q-1/p} \|u\|_p$,
- $L^p(\Omega) \hookrightarrow L^q(\Omega)$.

Corollary 2.10 (Interpolation in p) *Assume $\infty \geq p_1 \geq p \geq p_2 \geq 1$ and $u \in L^{p_1}(\Omega) \cap L^{p_2}(\Omega)$. Then*

$$\|u\|_p \leq \|u\|_{p_1}^{\alpha} \|u\|_{p_2}^{1-\alpha} \,,$$

where $\frac{1}{p} = \frac{\alpha}{p_1} + \frac{1-\alpha}{p_2}$, $\alpha \in [0,1]$.

Proof : Using Hölder's inequality we have

$$\|u\|_p = \left\{ \int_\Omega |u(x)|^{p\alpha} |u(x)|^{(1-\alpha)p} \, dx \right\}^{1/p} \leq \|u\|_{\alpha p \delta}^{\alpha} \|u\|_{(1-\alpha)p\delta'}^{1-\alpha} \,,$$

where $\frac{1}{\delta} + \frac{1}{\delta'} = 1$. The requirements $\alpha p \delta = p_1$ and $(1-\alpha)p\delta' = p_2$ imply $\frac{(1-\alpha)p}{p_2} + \frac{\alpha p}{p_1} = 1$, and the assertion follows. ■

Lemma 2.11 (Vitali) *Let Ω be a bounded domain in $\mathbb{R}^d$ and $f^n : \Omega \to \mathbb{R}$ be integrable for every $n \in \mathbb{N}$. Assume that*

- *$\lim_{n\to\infty} f^n(y)$ exists and is finite for almost all $y \in \Omega$;*
- *for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\sup_{n \in \mathbb{N}} \int_H |f^n(y)| \, dy < \varepsilon \qquad \forall\, H \subset \Omega,\ |H| < \delta \,.$$

Then

$$\lim_{n\to\infty} \int_\Omega f^n(y) \, dy = \int_\Omega \lim_{n\to\infty} f^n(y) \, dy \,.$$

Proof : See Alt [1992, p. 63] or Dunford and Schwartz [1958]. ■

1.2.4 Sobolev spaces

By $W^{k,p}(\Omega)$, $k \in \mathbb{N}$, $1 \le p \le \infty$, we mean the **Sobolev space** of all functions $u : \Omega \to \mathbb{R}$ having all distributional derivatives up to order k in $L^p(\Omega)$. The space $W^{k,p}(\Omega)$, equipped with the norm

$$\|u\|_{k,p} = \|u\|_{k,p;\Omega} \equiv \Big\{ \sum_{|\alpha| \le k} \|D^\alpha u\|_p^p \Big\}^{1/p} ,$$

is a Banach space. We identify $W^{0,p}(\Omega)$ with $L^p(\Omega)$. For k non-integer, we denote by $[k]$ the integer part of k. Then $W^{k,p}(\Omega)$ is a subspace of $W^{[k],p}(\Omega)$ consisting of functions $u \in W^{[k],p}(\Omega)$ for which

$$[D^\alpha u]^p_{k-[k],p} \equiv \int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{d+p(k-[k])}} dx\, dy$$

is finite for all α, $|\alpha| = [k]$. Then $W^{k,p}(\Omega)$ is a Banach space with the norm

$$\|u\|_{k,p} = \|u\|_{k,p;\Omega} \equiv \Big\{ \|u\|^p_{[k],p} + [D^\alpha u]^p_{k-[k],p} \Big\}^{1/p} , \quad k \notin \mathbb{N}.$$

For details see Nečas [1967], Wloka [1987] or Triebel [1978, 1992].

If no assumptions on Ω are made, it is understood that we consider two types of domains: either

$$\Omega \subset \mathbb{R}^d \quad \text{is open and bounded with } \partial\Omega \in C^{0,1} \tag{2.12}$$

or

$$\Omega = \mathbb{R}^d . \tag{2.13}$$

The space $W_0^{k,p}(\Omega)$ is defined as the closure of $\mathcal{D}(\Omega)$ with respect to the $W^{k,p}(\Omega)$-norm. The dual space of $W_0^{k,p}(\Omega)$ is denoted by $W^{-k,q}(\Omega)$, where $\frac{1}{q} + \frac{1}{p} = 1$.

Let us recall, see Nečas [1967], that for Ω satisfying (2.12) there exists a linear continuous operator $\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ called a **trace operator** such that

$$\begin{aligned} \gamma(u) = u|_{\partial\Omega} \qquad &\text{for } u \in C^1(\overline{\Omega}) , \\ \operatorname{Ker} \gamma = W_0^{1,p}(\Omega) , & \\ \operatorname{Range} \gamma = W_0^{1-\frac{1}{p},p}(\partial\Omega) \subset L^p(\partial\Omega) & \end{aligned} \tag{2.14}$$

and there exists $C > 0$ such that for all $u \in W^{1,p}(\Omega)$

$$\|\gamma(u)\|_{L^p(\partial\Omega)} \le \|\gamma(u)\|_{1-\frac{1}{p},p;\partial\Omega} \le C\|u\|_{1,p} \,. \tag{2.15}$$

If $\Omega = (0, L)^d$, then

$$W^{k,p}_{\text{per}}(\Omega) \equiv \overline{\{u \in C^\infty_{\text{per}}(\Omega), \int_\Omega u(x)\, dx = 0\}}^{W^{k,p}(\Omega)}$$

denotes the **Sobolev space of periodic functions**.

Lemma 2.16 *Assume $1 < q < \infty$, $k \in \mathbb{N}$ and $G \in W^{-k,q}(\Omega)$. Then there exist functions $\{g_\alpha\}_{|\alpha|\le k} \subset L^q(\Omega)$ such that*

$$G = \sum_{|\alpha|\le k} (-1)^{|\alpha|} D^\alpha g_\alpha \,.$$

Proof : See Nečas [1967, Chapter 2] or Adams [1975]. Let us note that g_α are in general not uniquely determined. ■

Theorem 2.17 (Imbeddings) *Let Ω satisfy (2.12) and let $0 \le j < k$, $1 \le p, q < \infty$. Put*

$$m_0 \equiv \frac{1}{p} - \frac{k-j}{d} \qquad \text{and} \qquad m \equiv \frac{1}{m_0} \quad \text{if} \quad m_0 \ne 0 \,.$$

- *Assume $m_0 > 0$. Then*

$$\begin{aligned} W^{k,p}(\Omega) &\hookrightarrow W^{j,m}(\Omega)\,, \\ W^{k,p}(\Omega) &\hookrightarrow\hookrightarrow W^{j,m_1}(\Omega)\,, \qquad m_1 < m\,, \\ W^{k,p}(\mathbb{R}^d) &\hookrightarrow W^{j,m}(\mathbb{R}^d)\,. \end{aligned}$$

- *Assume $m_0 < 0$. Then for $\alpha \in [0, 1)$*

$$\begin{aligned} m_0 + \frac{\alpha}{d} = 0 &\Rightarrow \begin{cases} W^{k,p}(\Omega) \hookrightarrow C^{j,\alpha}(\overline{\Omega})\,, \\ W^{k,p}(\mathbb{R}^d) \hookrightarrow C^{j,\alpha}(\mathbb{R}^d)\,, \end{cases} \\ m_0 + \frac{\alpha}{d} < 0 &\Rightarrow \quad W^{k,p}(\Omega) \hookrightarrow\hookrightarrow C^{j,\alpha}(\overline{\Omega})\,. \end{aligned}$$

- *Assume $m_0 = 0$. Then*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow W^{j,q}(\Omega)\,, \qquad q \in [1, \infty)\,.$$

Proof : If Ω satisfies (2.12) then the proofs can be found for example in Nečas [1967, Chapter 2] or Kufner, John and Fučík

[1977, Chapter 5]; if $\Omega = \mathbb{R}^d$ then the imbeddings are proved in NIKOLSKIJ [1975, Chapter 9] or ADAMS [1975], BERGH AND LÖFSTRÖM [1976], TRIEBEL [1978]. ■

Lemma 2.18 (Interpolation in k) *Let $k_1 \geq k_2 > 0$ with k_1, k_2 not necessarily integer. Then there exists a constant c such that for all $u \in W^{k_1,p}(\Omega)$*

$$\|u\|_{k_2,p} \leq c\|u\|_{k_1,p}^{\frac{k_2}{k_1}} \|u\|_p^{1-\frac{k_2}{k_1}} \quad . \tag{2.19}$$

PROOF : See for example ADAMS [1975, Chapter 7], BERGH AND LÖFSTRÖM [1976]. An elementary proof for $p = 2$ can be done by using Fourier transform, see for example ZEIDLER [1990a]. ■

Lemma 2.20 (Green's theorem) *Let $\Omega \subseteq \mathbb{R}^d$ satisfy (2.12) and $\mathbf{n} \equiv (n_1, \ldots, n_d)$ be the outward normal vector. Then for $u \in W^{1,1}(\Omega)$ we have*

$$\int_\Omega \frac{\partial u(x)}{\partial x_i}\, dx = \int_{\partial\Omega} u\, n_i\, ds\,, \qquad i = 1, \ldots, d\,,$$

where the values of u on $\partial\Omega$ are understood in the sense of traces.

PROOF : See for example NEČAS [1967, Section 3.1]. ■

1.2.5 Orlicz spaces

Let $\varphi : [0, \infty) \to \mathbb{R}$ be a non-negative nondecreasing right continuous function satisfying $\varphi(0) = 0$, $\varphi(\infty) \equiv \lim_{s\to\infty} \varphi(s) = \infty$. Then the function

$$\Phi(t) \equiv \int_0^{|t|} \varphi(s)\, ds\,, \tag{2.21}$$

usually called the **Young function** corresponding to φ, is even, continuous, convex, and satisfies

$$\lim_{t\to 0} \frac{\Phi(t)}{t} = 0\,, \qquad \lim_{t\to\infty} \frac{\Phi(t)}{t} = \infty\,.$$

Define $\psi(t) \equiv \sup_{\varphi(s)\leq t} s$. Let Ψ be the Young function corresponding to ψ. Then the functions Φ and Ψ are called **complementary Young functions**.

Let Φ and Ψ be complementary Young functions. Let us give some basic definitions:

1. The **Orlicz class** $\tilde{L}_\Phi(\Omega)$ consists of all measurable functions $u : \Omega \to \mathbb{R}$ such that
$$d(\Phi; u) \equiv \int_\Omega \Phi\big(|u(x)|\big)\, dx \tag{2.22}$$
is finite. It is known that in general $\tilde{L}_\Phi(\Omega)$ is not a linear space (see KUFNER, JOHN AND FUČÍK [1977, Section 3.1]).
2. Let us call the number
$$\|u\|_{L_\Phi(\Omega)} = \|u\|_\Phi \equiv \sup_{\substack{v \in \tilde{L}_\Psi(\Omega) \\ d(\Psi; v) \le 1}} \int_\Omega |u(x)v(x)|\, dx \tag{2.23}$$
the **Orlicz norm** of $u : \Omega \to \mathbb{R}$. The **Orlicz space** $L_\Phi(\Omega)$ is defined as the set of all measurable functions $u : \Omega \to \mathbb{R}$ for which the Orlicz norm (2.23) is finite.
3. By $C_\Phi(\Omega)$, we denote the closure of all bounded measurable compactly supported functions $u : \Omega \to \mathbb{R}$, with respect to the norm $\|\cdot\|_\Phi$. If $u \in C_\Phi(\Omega)$ then $\|u\|_{C_\Phi(\Omega)} \equiv \|u\|_{L_\Phi(\Omega)}$.

Example 2.24 The functions

1. t^p/p and t^q/q,
2. $\Phi(t) = e^t - t - 1$ and $\Psi(t) = (1+t)\log(1+t) - t$,
3. $\Phi_2(t) = e^{t^2} - 1$ and $\Psi_{1/2}$

are pairs of complementary Young functions. The explicit expression for function $\Psi_{1/2}$ is not known, nevertheless its asymptotic behaviour can be shown to be $t(\ln t)^{\frac{1}{2}}$ (KUFNER, JOHN AND FUČÍK [1977, Chapter 3]).

The Orlicz norm (2.23) requires knowledge of the complementary function. Sometimes, it is more convenient to use the so-called **Luxemburg norm** defined by
$$|||u|||_\Phi \equiv \inf\left\{\xi > 0\ ;\ \int_\Omega \Phi\left(\frac{|u(x)|}{\xi}\right) dx \le 1\right\}. \tag{2.25}$$
Both norms are equivalent (KUFNER, JOHN AND FUČÍK [1977, Section 3.8]). More precisely, for all $u \in L_\Phi(\Omega)$ we have
$$|||u|||_\Phi \le \|u\|_\Phi \le 2|||u|||_\Phi\ .$$

Lemma 2.26 *Let Φ, Ψ be two complementary Young functions. Then*

- *for all $u \in \tilde{L}_\Phi(\Omega)$, $v \in \tilde{L}_\Psi(\Omega)$*
$$\|u\,v\|_1 \le d(\Phi;u) + d(\Psi;v)\,; \tag{2.27}$$

- *for all $v \in L_\Psi(\Omega)$*
$$\|v\|_\Psi \le 1 + \int_\Omega \Psi(|v(x)|)\,dx\,; \tag{2.28}$$

- **(Hölder's inequality)** *for all $u \in L_\Phi(\Omega)$, $v \in L_\Psi(\Omega)$*
$$\|u\,v\|_1 \le \|u\|_\Phi \|v\|_\Psi\,. \tag{2.29}$$

PROOF : See KUFNER, JOHN AND FUČÍK [1977, Chapter 3]; note that the second assertion is a direct consequence of the first one. ■

Lemma 2.30 *Let Φ be a Young function. Then*

- *$L_\Phi(\Omega)$ is a Banach space;*
- *$C_\Phi(\Omega)$ is a separable Banach space and $\mathcal{D}(\Omega)$ is dense in $C_\Phi(\Omega)$;*
- *$C_\Phi(\Omega) \hookrightarrow \tilde{L}_\Phi(\Omega) \hookrightarrow L_\Phi(\Omega)$;*
- *$\left[C_\Phi(\Omega)\right]^* = L_\Psi(\Omega)$ provided that Ψ, Φ are complementary Young functions.*

PROOF : See KUFNER, JOHN AND FUČÍK [1977, Chapter 3]. ■

A Young function Φ satisfies the **Δ_2-condition** (in brief $\Phi \in \Delta_2$) if and only if there exist $c > 0$ and $t_0 \ge 0$ such that

$$\Phi(2t) \le c\,\Phi(t)$$

for every $t > t_0$. If Ω is unbounded, we require $t_0 = 0$. The following lemma shows the importance of Young functions satisfying the Δ_2-condition (cf. Lemma 2.30).

Lemma 2.31 *Let $\Phi \in \Delta_2$. Then*

- *$C_\Phi(\Omega) = \tilde{L}_\Phi(\Omega) = L_\Phi(\Omega)$;*
- *$\left[L_\Phi(\Omega)\right]^* = L_\Psi(\Omega)$, where Ψ is the complementary Young function to Φ;*
- *$L_\Phi(\Omega)$ is separable.*

PROOF : See KUFNER, JOHN AND FUČÍK [1977, Chapter 3]. ■

Let Φ_1, Φ_2 be two Young functions. We introduce the following orderings:

- $\Phi_1 < \Phi_2$ if and only if there exist $c > 0$ and $t_0 \geq 0$ such that $\Phi_1(t) \leq \Phi_2(ct)$ for all $t \geq t_0$. If Ω is unbounded, we require $t_0 = 0$.
- $\Phi_1 \ll \Phi_2$ if and only if $\lim\limits_{t\to\infty} \frac{\Phi_1(t)}{\Phi_2(\lambda t)} = 0$ for all $\lambda > 0$.

Lemma 2.32 *Let Φ_i, Ψ_i, $i = 1, 2$, be two pairs of complementary Young functions. If $\Phi_1 \ll \Phi_2$ then $\Psi_2 \ll \Psi_1$. If $\Phi_1 < \Phi_2$ then $\Psi_2 < \Psi_1$.*

PROOF : See KUFNER, JOHN AND FUČÍK [1977, Chapter 3]. ■

Lemma 2.33 *Let Φ_1, Φ_2 be two Young functions. Then*

- $\tilde{L}_{\Phi_2}(\Omega) \subset \tilde{L}_{\Phi_1}(\Omega)$ *if and only if* $\Phi_1 < \Phi_2$;
- $L_{\Phi_2}(\Omega) \hookrightarrow L_{\Phi_1}(\Omega)$ *if and only if* $\Phi_1 < \Phi_2$;
- *If* $\Phi_1 \ll \Phi_2$ *then* $L_{\Phi_2}(\Omega) \hookrightarrow C_{\Phi_1}(\Omega)$.

PROOF : See KUFNER, JOHN AND FUČÍK [1977, Chapter 3]. ■

Remark 2.34 Let $\Omega \subset \mathbb{R}^d$ satisfy (2.12) and $k, p \in \mathbb{R}$ be such that $kp > d$. Then due to Lemma 2.17

$$W^{k,p}(\Omega) \hookrightarrow C(\overline{\Omega}) \ . \tag{2.35}$$

Since $C(\overline{\Omega})$ is a subset of the set of all bounded measurable functions on $\overline{\Omega}$, we obviously have

$$W^{k,p}(\Omega) \subset C_\Phi(\Omega) \tag{2.36}$$

for any Young function Φ. Further, for $u \in W^{k,p}(\Omega)$

$$\begin{aligned}\|u\|_\Phi = \sup_{d(\Psi;v)\leq 1} \int_\Omega |u(x)v(x)|\, dx &\overset{(2.28)}{\leq} \|u\|_{C(\overline{\Omega})}\big[1 + \Phi(1)\mu(\Omega)\big] \\ &\equiv K\, \|u\|_{C(\overline{\Omega})} \ .\end{aligned}$$

By (2.35),

$$\|u\|_\Phi \leq c\, \|u\|_{k,p} \qquad \forall\, u \in W^{k,p}(\Omega)\, .$$

Thus, for any Young function Φ,

$$W^{k,p}(\Omega) \hookrightarrow C_\Phi(\Omega) \qquad \text{if } \ kp > d\, . \tag{2.37}$$

We will use this imbedding in Section 5.5.

Theorem 2.38 *Let $\Omega \subset \mathbb{R}^d$ satisfy (2.12). Let $kp = d$. Then*

$$W^{k,p}(\Omega) \hookrightarrow L_{\hat{\Phi}}(\Omega)$$

where $\hat{\Phi}(t) = \exp(|t|^{d/(d-1)}) - 1$. If $\Phi \ll \hat{\Phi}$ then the imbedding $W^{k,p}(\Omega)$ into $L_\Phi(\Omega)$ is compact.

PROOF : See for example KUFNER, JOHN AND FUČÍK [1977, Section 5.7] or TRUDINGER [1967]. ∎

Example 2.39 The Young functions Ψ and $\Psi_{1/2}$ defined in Example 2.24 satisfy the Δ_2-condition. Thus (see Lemma 2.31)

$$L_{\Psi_{1/2}}(\Omega) = C_{\Psi_{1/2}}(\Omega) \ , \quad L_\Psi(\Omega) = C_\Psi(\Omega) \ ,$$

and

$$\Big[L_\Psi(\Omega)\Big]^* = L_\Phi(\Omega) \ , \quad \Big[L_{\Psi_{1/2}}(\Omega)\Big]^* = L_{\Phi_2}(\Omega) \ .$$

Because $\Phi \ll \Phi_2$, we have $L_{\Phi_2}(\Omega) \hookrightarrow L_\Phi(\Omega)$. Moreover, from Lemma 2.30 we obtain $C_{\Phi_2}(\Omega) \hookrightarrow C_\Phi(\Omega)$ and due to Lemma 2.32 we also have $L_\Psi(\Omega) \hookrightarrow L_{\Psi_{1/2}}(\Omega)$.

Theorem 2.40 (Absolute continuity of the Orlicz norm) *Let $\Phi \in \Delta_2$ and $g \in L_\Phi(\Omega)$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|g\|_{L_\Phi(\Omega')} \le \varepsilon$ provided that $|\Omega'| < \delta$, $\Omega' \subset \Omega$.*

PROOF : See KRASNOSELSKII AND RUTICKII [1958, Chapter 2] or RAO AND REN [1991]. ∎

1.2.6 Bochner spaces

The main part of this book is devoted to evolution problems. In this case, one of the variables—time t—can be understood in a different way to the other ones, mostly denoted by $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d)$.

Let $\Omega \subset \mathbb{R}^d$ and $T > 0$. By Q_T we denote a time-space cylinder $Q_T \equiv I \times \Omega$, where $I \equiv (0, T)$. For $u : Q_T \to \mathbb{R}$ the map

$$u(t) : x \longmapsto u(t, x)$$

is an element of some function space (for instance Lebesgue, Hölder, Sobolev, Orlicz space). Then the function

$$t \longmapsto u(t)$$

maps the time interval I into that function space. Spaces of functions which map time interval I into some Banach space are called **Bochner spaces**.

Let X be a Banach space. The space $C^k(I;X)$, $k \in \mathbb{N}$, contains all continuous functions $u : \overline{I} \to X$ for which all (time-)derivatives, up to order k can be continuously extended to $\overline{I}$. In particular,

$$\|u\|_{C^k(I;X)} \equiv \sum_{s=0}^{k} \sup_{[0,T)} \Big\| \frac{\partial^s u(t)}{\partial t^s} \Big\|_X$$

is finite and $\|\cdot\|_{C^k(I;X)}$ defines the norm in $C^k(I;X)$.

Similarly, by $L^p(I;X)$, $1 \le p \le \infty$, we denote the space of all measurable functions¶ $u : I \to X$ for which the norm

$$\|u\|_{L^p(I;X)} \equiv \Big\{ \int_0^T \|u(t)\|_X^p \, dt \Big\}^{\frac{1}{p}}, \qquad p < \infty$$

or

$$\|u\|_{L^\infty(I;X)} \equiv \operatorname*{ess\,sup}_{t \in I} \|u(t)\|_X$$

respectively, is finite.

Let us summarize some basic properties of Bochner spaces.

Lemma 2.41 (Hölder's inequality) *Let X be a Banach space and $\frac{1}{p} + \frac{1}{q} = 1$. Assume $u \in L^p(I;X)$, $v \in L^q(I;X^*)$. Then $\langle v(\cdot), u(\cdot)\rangle_X \in L^1(I)$ and*

$$\int_0^T |\langle v(s), u(s)\rangle_X| \, ds \le \|u\|_{L^p(I;X)} \, \|v\|_{L^q(I;X^*)} \, .$$

PROOF : See ZEIDLER [1990b, Chapter 23]. ■

Lemma 2.42 *Let X, Y be Banach spaces. Then*

- *$X \hookrightarrow Y$ implies $L^q(I;X) \hookrightarrow L^p(I;Y)$ if $1 \le p \le q \le \infty$;*
- *if $1 \le p < \infty$ then $C(I;X)$ is dense in $L^p(I;X)$ and*

$$C(I;X) \hookrightarrow L^p(I;X) \, .$$

PROOF : See ZEIDLER [1990b, Chapter 23]. ■

¶ A function $u : I \to X$ is (strongly) measurable if and only if there exist step functions $u_n : I \to X$ such that $\|u_n(t) - u(t)\|_X \to 0$ for almost all $t \in I$.

Now let us clarify in what way we understand the time derivative of a function $u : I \to X$ from a Bochner space. We consider the following situation:

Let H be a Hilbert space with a scalar product $(\cdot,\cdot)_H$ and let X be a Banach space such that

$$X \hookrightarrow H \simeq H^* \hookrightarrow X^* \tag{2.43}$$

and

$$X \text{ is dense in } H\,. \tag{2.44}$$

Then, if $u \in L^p(I;X)$, we denote by $\frac{du}{dt}$ an element of the space $L^{p'}(I;X^*)$ (where $1/p + 1/p' = 1$) such that

$$\int_0^T \Big\langle \frac{du(t)}{dt}, v \Big\rangle_X \varphi(t)\,dt = -\int_0^T (u(t),v)_H \varphi'(t)\,dt$$

for all $v \in X$ and $\varphi \in \mathcal{D}(I)$. Note that the scalar product $(\cdot,\cdot)_H$ in the above formula must be appropriately chosen. For example, for $W^{2,2}(\Omega) \hookrightarrow W^{1,2}(\Omega) \hookrightarrow \big(W^{2,2}(\Omega)\big)^*$ the appropriate one is

$$(u,v)_H = (u,v)_{L^2} + (\nabla u, \nabla v)_{L^2}\,.$$

Lemma 2.45 *Let (2.43), (2.44) be satisfied and let $p \in (1,\infty)$. Then*

- $W \equiv \{u \in L^p(I;X); \frac{du}{dt} \in L^{p'}(I;X^*)\} \hookrightarrow C(I;H)$*;*
- **(Partial integration)** *for all $u, v \in W$ and all $s, t \in I$*

$$\begin{aligned}(u(t),v(t))_H &- (u(s),v(s))_H \\ &= \int_s^t \Big\langle \frac{du(\tau)}{dt}, v(\tau)\Big\rangle_X + \Big\langle \frac{dv(\tau)}{dt}, u(\tau)\Big\rangle_X \,d\tau\,.\end{aligned} \tag{2.46}$$

In particular for $u = v$,

$$\frac{1}{2}\|u(t)\|_H^2 - \frac{1}{2}\|u(s)\|_H^2 = \int_s^t \Big\langle \frac{du(\tau)}{dt}, u(\tau)\Big\rangle_X \,d\tau \tag{2.47}$$

for all $s, t \in I$.

Proof : See Gajewski, Gröger and Zacharias [1974, Section 4.1]. ∎

The following lemma on the compact imbedding in Bochner spaces is very important:

Lemma 2.48 (Aubin-Lions) *Let $1 < \alpha, \beta < +\infty$. Let X be a Banach space, and let X_0, X_1 be separable and reflexive Banach spaces. Provided that $X_0 \hookrightarrow\hookrightarrow X \hookrightarrow X_1$ we have*

$$\left\{ v \in L^\alpha(I; X_0);\ \frac{dv}{dt} \in L^\beta(I; X_1) \right\} \hookrightarrow\hookrightarrow L^\alpha(I; X)\,.$$

PROOF : See LIONS [1969, Section 1.5] or the survey paper SIMON [1987]. A generalized form of this lemma for locally convex spaces and $\beta = 1$ can be found in ROUBÍČEK [1990]. ■

1.2.7 The space of functions with bounded variation

Let $u \in L^1(\Omega)$. The **total variation** of u is defined as

$$TV_\Omega(u) \equiv \sup_{\substack{\varphi \in \mathcal{D}(\Omega)^d \\ \|\varphi\|_\infty \le 1}} \int_\Omega u \, \operatorname{div} \varphi \, dx\,.$$

By $BV(\Omega)$ we denote the subspace of functions $u \in L^1(\Omega)$ with $TV_\Omega(u) < \infty$,

$$BV(\Omega) \equiv \left\{ u \in L^1(\Omega); TV_\Omega(u) < \infty \right\}.$$

Let $u \in BV(\Omega)$ and put

$$\|u\|_{BV(\Omega)} \equiv \|u\|_1 + TV_\Omega(u)\,. \tag{2.49}$$

Lemma 2.50

- *$\|\cdot\|_{BV(\Omega)}$ is a norm on $BV(\Omega)$, and $BV(\Omega)$ with $\|\cdot\|_{BV(\Omega)}$ is a Banach space.*
- *$W^{1,1}(\Omega) \hookrightarrow BV(\Omega) \hookrightarrow L^1(\Omega)$.*

PROOF : It is not difficult to see that $\|\cdot\|_{BV(\Omega)}$ is a norm. The following lower semicontinuity property plays the key role in the proof of completeness of $BV(\Omega)$:

$$\left[u_n \to u \quad \text{in } L^1(\Omega) \right] \Longrightarrow TV_\Omega(u) \le \liminf_{n\to\infty} TV_\Omega(u_n)\,. \tag{2.51}$$

Assuming (2.51) we can prove that $BV(\Omega)$ is complete. Indeed, if u_n is a Cauchy sequence in $BV(\Omega)$, then there exists some $u \in L^1(\Omega)$ such that

$$u_n \to u \qquad \text{in } L^1(\Omega)\,. \tag{2.52}$$

According to (2.51), $TV_\Omega(u) < \infty$ and $u \in BV(\Omega)$. It remains to prove

$$TV_\Omega(u_n - u) \overset{n\to\infty}{\to} 0\,.$$

However, for all $\varepsilon > 0$

$$TV_\Omega(u_n - u_k) < \varepsilon \qquad \text{for } \ n,k \quad \text{large enough.}$$

Using (2.52) we get

$$(u_n - u_k) \stackrel{k\to\infty}{\to} (u_n - u) \qquad \text{in } \ L^1(\Omega)\,.$$

This together with (2.51) gives

$$TV_\Omega(u_n - u) \le \liminf_{k\to\infty} TV_\Omega(u_n - u_k) \le \varepsilon \quad \text{for } \ n \ \text{ large enough,}$$

which proves the completeness of $BV(\Omega)$.

It remains to verify the validity of (2.51). Let $\boldsymbol{\varphi} \in \mathcal{D}(\Omega)^d$ with $\|\boldsymbol{\varphi}\|_\infty \le 1$. It follows from the assumption $u_n \to u$ in $L^1(\Omega)$ that

$$\lim_{n\to\infty} \int_\Omega u_n \operatorname{div} \boldsymbol{\varphi}\, dx = \int_\Omega u \operatorname{div} \boldsymbol{\varphi}\, dx\,. \tag{2.53}$$

But for each $n \in \mathbb{N}$

$$\int_\Omega u_n \operatorname{div} \boldsymbol{\varphi}\, dx \le TV_\Omega(u_n)\,.$$

Passing to the $\liminf$ in the previous inequality as $n \to \infty$ and using (2.53) we get the assertion.

The second statement follows from the definition of the corresponding spaces. ∎

Lemma 2.54 (Imbeddings) *Let Ω satisfy (2.12). If $d > 1$ then*

$$BV(\Omega) \hookrightarrow L^{\frac{d}{d-1}}(\Omega)\,,$$
$$BV(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega)\,, \quad \forall\, 1 \le q < \frac{d}{d-1}\,.$$

PROOF : See FEDERER [1969] or GIUSTI [1984, Chapter 1]. ∎

1.2.8 Radon measures

Let Ω be a bounded domain. We denote by $M(\Omega)$ the space of the so-called **Radon measures** defined as the dual space of $C(\overline{\Omega})$.

Clearly $L^1(\Omega) \hookrightarrow M(\Omega)$ since for $f \in L^1(\Omega)$

$$F(\varphi) \equiv \int_\Omega f(x)\varphi(x)\, dx \qquad \forall\, \varphi \in C(\overline{\Omega})$$

defines a continuous linear functional on $C(\overline{\Omega})$ and consequently $F \in M(\Omega)$. Moreover,

$$\|F\|_{M(\Omega)} \le \|f\|_1\,,$$

where $\|\cdot\|_{M(\Omega)}$ is defined as a dual norm. Thus, the space of (Radon) measures represents a natural extension of the space of integrable functions.

Lemma 2.55 (Imbeddings) *Let Ω satisfy (2.12) and let $\{\nu_k\} \subset M(\Omega)$ be a bounded sequence in $M(\Omega)$. Then for every $1 \le q < \frac{d}{d-1}$ there exist $\nu \in W^{-1,q}(\Omega)$ and a subsequence $\{\nu_{k_j}\} \subset \{\nu_k\}$ such that*

$$\nu_{k_j} \to \nu \qquad \text{in } W^{-1,q}(\Omega) \quad \text{as } j \to \infty\,.$$

PROOF : (See also EVANS [1990]). Since $C(\overline{\Omega})$ is separable, it follows from Theorem 2.1 that there exist $\nu \in M(\Omega)$ and $\{\nu_{k_j}\} \subset \{\nu_k\}$ such that

$$\nu_{k_j} \overset{*}{\rightharpoonup} \nu \qquad \text{in } M(\Omega)\,,$$

i.e.,

$$\int_\Omega \Phi \, d\nu_{k_j} \to \int_\Omega \Phi \, d\nu \qquad \forall\, \Phi \in C(\overline{\Omega}) \text{ as } j \to \infty\,. \tag{2.56}$$

Let $1 \le q < \frac{d}{d-1}$ and let q' be the dual index to q, i.e. $\frac{1}{q} + \frac{1}{q'} = 1$. Our aim is to show

$$\sup_{\substack{\Phi \in W_0^{1,q'}(\Omega) \\ \|\Phi\|_{1,q'} \le 1}} \Big| \int_\Omega \Phi \, d\nu_{k_j} - \int_\Omega \Phi \, d\nu \Big| \to 0 \qquad \text{as } j \to \infty\,.$$

Since $q' > d$, Theorem 2.17 implies that the unit ball in $W_0^{1,q'}(\Omega)$ is precompact in $C(\overline{\Omega})$. Consequently, for every $\Phi \in W_0^{1,q'}(\Omega)$, $\|\Phi\|_{1,q'} \le 1$, and for all $\varepsilon > 0$ there exists a net $\{\Phi_i\}_{i=1}^{N(\varepsilon)} \subset W_0^{1,q'}(\Omega)$ such that

$$\min_{1 \le i \le N(\varepsilon)} \|\Phi - \Phi_i\|_\infty \le \varepsilon\,.$$

Thus the following calculation finishes the proof:

$$\begin{aligned} &\Big| \int_\Omega \Phi \, d\nu_{k_j} - \int_\Omega \Phi \, d\nu \Big| \\ &\qquad \le \Big| \int_\Omega (\Phi - \Phi_i) d\nu_{k_j} \Big| + \Big| \int_\Omega \Phi_i d\nu_{k_j} - \int_\Omega \Phi_i d\nu \Big| \\ &\qquad\qquad + \Big| \int_\Omega (\Phi_i - \Phi) d\nu \Big| \\ &\qquad \le 2\varepsilon \sup_j |\nu_{k_j}|(\Omega) + \Big| \int_\Omega \Phi_i d\nu_{k_j} - \int_\Omega \Phi_i d\nu \Big|\,. \end{aligned} \tag{2.57}$$

Due to (2.56), the last term in (2.57) tends to zero as $j \to \infty$. ■

If $\Omega = \mathbb{R}^d$ (a locally compact set which is not compact) we define

$$C_0(\mathbb{R}^d) \equiv \left\{ u \in C(\mathbb{R}^d);\ \lim_{|x|\to\infty} u(x) = 0 \right\}.$$

Note that

$$C_0(\mathbb{R}^d) = \overline{\mathcal{D}(\mathbb{R}^d)}^{\|\cdot\|_\infty}.$$

The space of Radon measures is defined as

$$\begin{aligned} M(\mathbb{R}^d) \equiv \big\{ \mu : C_0(\mathbb{R}^d) \to \mathbb{R}; \mu \text{ linear s. t. } \exists\, c > 0 \\ |\mu(f)| \le c\|f\|_\infty\ \forall\, f \in \mathcal{D}(\mathbb{R}^d) \big\}. \end{aligned} \tag{2.58}$$

Let us further define

$$\|\mu\|_{M(\mathbb{R}^d)} \equiv \sup_{\substack{f\in\mathcal{D}(\mathbb{R}^d) \\ \|f\|_\infty \le 1}} |\mu(f)|. \tag{2.59}$$

If $\mu \in M(\mathbb{R}^d)$, $\mu(f) \ge 0$ for all $f \in \mathcal{D}(\mathbb{R}^d)$, $f \ge 0$ we say that μ is a non-negative bounded Radon measure.

The space of **probability measures** is then defined as follows:

$$\mathrm{Prob}(\mathbb{R}^d) \equiv \{\mu \in M(\mathbb{R}^d),\ \mu \text{ non-negative},\ \|\mu\|_{M(\mathbb{R}^d)} = 1\}. \tag{2.60}$$

Lemma 2.61

- *The space* $(M(\mathbb{R}^d), \|\cdot\|_{M(\mathbb{R}^d)})$ *is a Banach space.*
- *Any* $\mu \in M(\mathbb{R}^d)$ *can be uniquely extended to the element of dual space to* $C_0(\mathbb{R}^d)$. *In this sense,*

$$\big(C_0(\mathbb{R}^d), \|\cdot\|_\infty\big)^* \equiv M(\mathbb{R}^d). \tag{2.62}$$

PROOF : See DUNFORD AND SCHWARTZ [1958], BOURBAKI [1965]. ■

Remark 2.63 The functionals $\mu \in M(\mathbb{R}^d)$ are called (Radon) measures, since there is a one-to-one correspondence between elements of $M(\mathbb{R}^d)$ and a class of (Borel) measures $\widetilde{\mu}$ on $\mathbb{R}^d$ with finite total mass $\widetilde{\mu}(\mathbb{R}^d) < \infty$, such that

$$\mu(f) = \int_{\mathbb{R}^k} f\, d\widetilde{\mu} \qquad \forall\, f \in C_0(\mathbb{R}^d). \tag{2.64}$$

As usual, we do not distinguish between μ and $\widetilde{\mu}$. Finally, instead of $\mu(f)$ we use the standard duality notation

$$\langle \mu, f \rangle \equiv \mu(f) = \int_{\mathbb{R}^d} f \, d\mu \,, \quad \mu \in M(\mathbb{R}^d), \; f \in C_0(\mathbb{R}^d) \,. \tag{2.65}$$

For more details see BOURBAKI [1965].