UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Part I

Introduction

Chapter 1

A Brief History of PDEs

1.1 From Newton to Schwarz

The advent of the Differential and Integral Calculus in the later half of the seventeenth century heralded the start of a new age in Mathematics. This is true both in regards to the applications of Mathematics to other sciences, notably to Physics, Economics, Chemistry and lately also Biology, and also in respect of the development of Mathematics as such. Indeed, in connection with the second aspect of these new developments, we may note that from the Differential and Integral Calculus, and the new point of view it introduced in so far as mathematical functions are concerned, the vast and powerful field of Mathematical Analysis developed. On the other hand, and parallel to the above mentioned development of abstract Mathematics, Newton's Calculus provides powerful new tools with which to solve so called "real world problems". Indeed, it was exactly the consideration of physical problems, namely, The Laws of Motion, which lead Newton to the conception of the Calculus.

With respect to the above mentioned power of the Calculus when it comes to the mathematical solution of physical problems, we need only note the following. Prior to the invention of the Differential Calculus, the only type of motion that could be described in a mathematically precise way was that of a particle moving uniformly along a straight line, and that of a particle moving at constant angular momentum along a circular path. In contradistinction with this rudimentary earlier state of affairs, and as is well known, Newton's Differential and Integral Calculus provides the appropriate mathematical machinery for the formulation, in precise mathematical terms, and solution of the basic laws of nature, thus going incomparably farther than the mentioned simple motions.

The mentioned mathematical expressions of these laws, using the Calculus, typically take the form of systems of ordinary differential equations (ODEs), or systems of partial differential equations (PDEs). At first, such systems of equations were mostly of a particular form, namely, linear and at most of second order. However, with the emergence of increasingly sophisticated physical theories, and the development of state of the art technologies, in particular during the second half of the twentieth century, the need arose for more complex mathematical models, which typically take the form of systems of *nonlinear* ODEs or PDEs. As such it is clear that the theoretical treatment of nonlinear PDEs, in particular in connection with the existence of the solutions to such equations, and the properties of these solutions, should such solutions exists, is of major interest.

In this regard, and for over a century by now, there have been general and type independent results on the existence and regularity of the solutions to both systems of ODEs and systems of PDEs. In particular, in the case of systems of ODEs, one may recall Picard's Theorem [125].

Theorem 1 **[125] Consider any system of K ODEs in K unknown functions $u_1, ..., u_K$ of the form*

$$\frac{d\boldsymbol{y}}{dt} = \boldsymbol{F}\left(t, y_1\left(t\right), ..., y_K\left(t\right)\right), \tag{1.1}$$

where $\boldsymbol{F} : \mathbb{R}\times\mathbb{R}^K \to \mathbb{R}^K$ is continuous on a neighborhood V of the point $(t_0, (y_{1,0}, ..., y_{K,0})) \in \mathbb{R} \times \mathbb{R}^K$. Then there is some $\delta > 0$ which depends continuously on the initial data $(t_0, (y_{1,0}, ..., y_{K,0}))$, and a solution $\boldsymbol{y} \in \mathcal{C}^1\left(t_0 - \delta, t_0 + \delta\right)^K$ to (1.1) that also satisfies the initial condition

$$\boldsymbol{y}\left(t_0\right) = \left(y_{i,0}\right).$$

Furthermore, if $\boldsymbol{F}$ is Lipschitz on V, then the solution is unique.

In modern accounts of the theory of ODEs, Theorem 1 is typically presented as an application of Banach's fixed point principle in Banach spaces. This might lead to the impression that this result is obtained as an application of functional analysis. However, and as mentioned, Picard's proof [125] of Theorem 1 predates the formulation of linear functional analysis, which culminated in Banach's similar work [17]. In fact, Picard's proof is based on techniques from the classical theory of functions, notably integration of usual smooth functions. One may also note that, around the same time that Picard proved Theorem 1, Peano [123] gave a proof of the *existence* of a solution of (1.1), which is based on the Arzellà-Ascoli Theorem.

In the case of systems of PDEs, the first comparable general and type independent existence and regularity result is due to Kovalevskaia [86]. It is interesting, in view of the common perception that ODEs are far simpler objects than PDEs, to note that the Cauchy-Kovalevskaia Theorem precedes Picard's Theorem by about twenty years, and as such is not, and in fact could not at the time, be based on more advanced mathematics.

Theorem 2 **[86] Consider the system of K nonlinear partial differential equations of the form*

$$D_t^m \boldsymbol{u}\left(t, y\right) = \boldsymbol{G}\left(t, y, ..., D_t^p D_y^q u_i\left(t, y\right), ...\right) \tag{1.2}$$

with $t \in \mathbb{R}$, $y \in \mathbb{R}^{n-1}$, $m \geq 1$, $0 \leq p < m$, $q \in \mathbb{N}^{n-1}$, $p + |q| \leq m$ *and with the analytic Cauchy data*

$$D_t^p \boldsymbol{u}(t_0, y) = \boldsymbol{g}_p(y), \; 0 \leq p < m, \; (t_0, y) \in S \tag{1.3}$$

on the noncharacteristic analytic hypersurface

$$S = \{(t_0, y) \; : y \in \mathbb{R}^{n-1}\}.$$

If the mapping $\boldsymbol{G}$ *is analytic, then there exists a neighborhood* V *of* t_0 *in* $\mathbb{R}$*, and an analytic function* $\boldsymbol{u} : V \times \mathbb{R}^{n-1} \to \mathbb{C}^K$ *that satisfies (1.2) and (1.3).*

It should be noted that the two general and type independent existence and regularity results, namely, Theorem 1 and Theorem 2, for systems of ODEs, respectively PDEs, predates the invention of functional analysis by nearly fifty years. Moreover, and as all ready mentioned in connection with Theorem 1, these results are based on some comparatively elementary mathematics. In particular, the only so called "hard mathematics" involved in the proof of Theorem 2 is the power series expansion for an analytic function, and certain Abel-type estimates for these expansions.

Both of the existence results, Theorem 1 and Theorem 2, are *local* in nature. That is, the solution cannot be guaranteed to exists on the *whole* domain of definition of the respective systems of equations. Furthermore, and as can be seen from rather simple examples, this is not due to the limitations of the particular techniques used to prove these results, but may instead be attributed to the very nature of nonlinear ODEs and PDEs in such a general setup.

The mentioned local nature of Theorems 1 and 2 is unsatisfactory in at least two respects. In the first place, in many of the physical problems that are supposed to be modeled by the respective system of ODEs or PDEs, we may be interested in solutions which exists on domains that are larger than that delivered by the respective existence results in Theorem 1 and Theorem 2, respectively. Secondly, and as can be seen from rather elementary examples, classical solutions, such as those obtained in the mentioned existence results, will in general fail to exists on the entire domain of physical interest. In this regard, a particularly simple, yet relevant, example is the conservation law

$$U_t + U_x U = 0, \; t > 0, \; x \in \mathbb{R} \tag{1.4}$$

with the initial condition

$$U(0, x) = u(x), \; x \in \mathbb{R}. \tag{1.5}$$

If we assume that the function u in (1.5) is smooth enough, then a classical solution U, in fact an analytic solution, of (1.4) to (1.5) is given by the implicit equation

$$U(t, x) = u(x - tU(t, x)), \; t \geq 0, \; x \in \mathbb{R}. \tag{1.6}$$

According to the implicit function theorem, we can obtain $U(s, y)$ from (1.6) for s and y in suitable neighborhoods of t and s, respectively, whenever

$$tu'(x - tU(t, x)) + 1 \neq 0. \tag{1.7}$$

For $t = 0$ the condition (1.7) is clearly satisfied. As such, there is a neighborhood $\Omega \subseteq [0, \infty) \times \mathbb{R}$ of the x-axis $\mathbb{R}$ so that $U(t, x)$ exists for $(t, x) \in \Omega$. However, if for some interval $I \subseteq \mathbb{R}$

$$u'(x) < 0,\ x \in I \tag{1.8}$$

then for certain values $t > 0$, the condition (1.7) may fail, irrespective of the domain or degree of smoothness of u. It is well known that the violation of the condition (1.7) may imply that the classical solution U fails to exists for the respective values of t and x. That is, the domain of existence Ω of the solution U will be strictly contained in $[0, \infty) \times \mathbb{R}$. As such, for certain $x \in \mathbb{R}$, the solution $U(t, x)$ does not exist for sufficiently large $t > 0$ so that the equation (1.4) fails to have a classical solution on the whole of its domain of definition.

From a physical point of view, however, it is exactly the points $(t, x) \in ([0, \infty) \times \mathbb{R}) \setminus \Omega$ where the solution fails to exist that are of interest, as these points may represent the appearance and propagation of what are called *shock waves*. Under rather general conditions, see for instance [96] and [143], it is possible to define certain *generalized solutions* U for all $t \geq 0$ and $x \in \mathbb{R}$, which turn out to be physically meaningful, and which are in fact classical solutions everywhere on $[0, \infty) \times \mathbb{R}$ except a suitable set of points $\Gamma \subset [0, \infty) \times \mathbb{R}$, where Γ consists of certain families of curves called *shock fronts*.

As a clarification of the above mentioned *lack of global smoothness* of the solutions to (1.4) and (1.5), let us consider the example

$$u(x) = \begin{cases} 1 & if \quad x \leq 0 \\ 1 - x & if \quad 0 \leq x \leq 1 \\ 0 & if \quad x \geq 1 \end{cases}. \tag{1.9}$$

In this case, the shock front Γ is given by

$$\Gamma = \left\{ (t, x) \left| \begin{array}{ll} 1) & t \geq 1 \\ 2) & x = \frac{t+1}{2} \end{array} \right. \right\}, \tag{1.10}$$

while the solution $U(t, x)$ is given by

$$U(t, x) = \begin{cases} 1 & if \quad x \leq 0 \\ \frac{x-1}{t-1} & if \quad 0 \leq x \leq 1 \\ 0 & if \quad x \geq 1 \end{cases} \tag{1.11}$$

when $0 \leq t \leq 1$ and

$$U(t,x) = \begin{cases} 1 & if \quad x < \frac{t+1}{2} \\ 0 & if \quad x > \frac{t+1}{2} \end{cases} \tag{1.12}$$

when $t \geq 1$. For $(t,x) \in \Gamma$ one may define $U(t,x)$ at will. In this example, the failure of the initial value u to be sufficiently smooth at $x = 0$ and $x = 1$ does not in any way contribute to the nonexistence of a solution $U(t,x)$ which is classical on the whole domain of definition of the equations. Rather, the lack of global smoothness of the solution $U(t,x)$ is due to the fact that the initial condition u satisfies (1.8) on the interval $(0,1) \subset \mathbb{R}$.

In view of the local nature of the existence results in Theorem 1 and Theorem 2, as well as the occurrence of singularities in the solutions of nonlinear PDEs as demonstrated in the above example concerning the nonlinear conservation law (1.4), it is clear that there is both a physical and theoretical interest in the existence of solutions to such systems of PDEs that may fail to be classical on the whole domain of definition of the respective system of PDEs.

The interest in such *generalized solutions* to PDEs is the main motivation for the study of generalized functions, that is, objects which retain certain essential features of usual real valued functions. In this regard, there have so far been two main approaches to constructing suitable spaces of generalized functions which may contain suitable generalized solutions to systems of PDEs, namely, the sequential approach and the functional analytic approach.

The sequential approach, introduced by S L Sobolev, see for instance [148] and [149], is based on the concept of a weak derivative, which had been applied by Hilbert, Courant, Riemann and several others in the study of various classes of ODEs and PDEs. In this regard, we recall that a Lebesgue measurable function $u : \Omega \rightarrow \mathbb{R}$ is square integrable on Ω whenever

$$\int_{\Omega} u(x)^2 \, dx < \infty \tag{1.13}$$

where the integral is taken in the sense of Lebesgue. The set of all square integrable functions on Ω is denoted $L_2(\Omega)$, and carries the structure of a Hilbert space under the inner product

$$\langle u, v \rangle_{L_2} = \int_{\Omega} v(x) \, u(x) \, dx.$$

For $u \in L_2(\Omega)$ and $\alpha \in \mathbb{N}^n$, a measurable function $v : \Omega \rightarrow \mathbb{R}$ is called a weak derivative $D^\alpha u$ of u whenever

$$\begin{array}{ll} \forall & K \subset \Omega \text{ compact :} \\ \forall & \varphi \in \mathcal{D}(K) \ : \\ & \int_K u(x) \, D^\alpha \varphi(x) \, dx = (-1)^{|\alpha|} \int_K v(x) \, \varphi(x) \, dx \end{array} \tag{1.14}$$

where $\mathcal{D}(K)$ is the set of $\mathcal{C}^{\infty}$-smooth functions φ on K such that the closure of the set

$$\{x \in K : \varphi(x) \neq 0\}$$

is compact and strictly contained in K, or some other suitable space of test functions.

In this sequential approach, generalized solutions to a nonlinear, or linear, PDE

$$T(x, D)\, u(x) = 0,\ x \in \Omega \subseteq \mathbb{R}^n \tag{1.15}$$

are obtained by constructing a sequence of approximating equations

$$T_i(x, D)\, u_i(x) = 0,\ x \in \Omega,\ i \in \mathbb{N} \tag{1.16}$$

where the operators $T_i(x, D)$ are supposed to approximate $T(x, D)$ in a prescribed way, so that each u_i is a classical solution of (1.16), and the sequence (u_i) converges in a suitable *weak* sense to a function u, for instance

$$\int_{\Omega} (u_i(x) - u(x))\, \varphi(x)\, dx \to 0 \tag{1.17}$$

for suitable test functions φ. The weak limit u of the sequence (u_i) is interpreted as a generalized solution of (1.15).

In this regard, Sobolev introduced the space $H^{2,m}(\Omega)$, with $m \geq 1$, which is defined as

$$H^{2,m}(\Omega) = \left\{ u \in L_2(\Omega) \middle| \begin{array}{ll} \forall & |\alpha| \leq m : \\ & D^{\alpha} f \in L_2(\Omega) \end{array} \right\}. \tag{1.18}$$

That is, the Sobolev space $H^{2,m}(\Omega)$ consists of all square integrable functions u on Ω with all weak partial derivatives $D^{\alpha}u$ up to order m in $L_2(\Omega)$. An inner product may be defined on $H^{2,m}(\Omega)$ through the formula

$$\langle u, v \rangle_m = \sum_{|\alpha| \leq m} \langle D^{\alpha}u, D^{\alpha}v \rangle_{L_2}. \tag{1.19}$$

so that the Sobolev space $H^{2,m}(\Omega)$ is a Hilbert space. In particular, a sequence (u_i) in $H^{2,m}(\Omega)$ converges to $u \in H^{2,m}(\Omega)$ if and only if

$$\sum_{|\alpha| \leq m} \int_{\Omega} (D^{\alpha}u(x) - D^{\alpha}u_i(x))^2\, dx \to 0$$

which, in particular, implies that the sequence (u_i) converges weakly to u, in the sense that

$$\begin{array}{ll} \forall & \varphi \in \mathcal{D}(\Omega) : \\ & \sum_{|\alpha| \leq m} \int_{\Omega} (u(x) - u_i(x))\, D^{\alpha}\varphi(x)\, dx \to 0 \end{array}$$

The definition of the Hilbert space structure on $H^{2,m}(\Omega)$ is the essential feature introduced by Sobolev. Indeed, the concept of a weak derivative, and that of weak solution, had been used by many authors prior to Sobolev. However, the main difficulty in applying the techniques of modern analysis, in particular those connected with function spaces and topological structures on such spaces, to ODEs and PDEs, respectively, is that the differential operators on such spaces are typically not continuous with respect to the mentioned topological structures. In the case of the Sobolev spaces $H^{2,m}(\Omega)$, and in view of (1.19), it is clear that

$$\begin{array}{ll} \forall & |\alpha| \leq m : \\ \forall & u \in H^{2,m}(\Omega) : \\ & \langle D^\alpha u, D^\alpha u\rangle_{L_2} \leq \langle u, u\rangle_m \end{array}$$

so that each such differential operator

$$D^\alpha : H^{2,m}(\Omega) \rightarrow L_2(\Omega) \tag{1.20}$$

is continuous with respect to the inner products on $H^{2,m}(\Omega)$ and $L_2(\Omega)$, respectively. As such, the powerful tools of analysis, and in particular linear functional analysis, may be applied to the study of such PDEs which admit a suitable weak formulation in terms of $H^{2,m}(\Omega)$ or other associated spaces. At this point it is worth noting that this sequential approach has as of yet not received any suitable general theoretic treatment. Nevertheless, it has resulted in a wide range of effective, though somewhat ad hoc, solution methods for both linear and nonlinear PDEs, see for instance [100].

The second major approach to establishing generalized solutions to PDEs within a suitable framework of generalized functions, namely, the functional analytic approach introduced by L Schwartz [144], [145] in the late 1940s is based on the the idea of generalizing the concept of a weak derivative through the machinery of linear function analysis.

In this regard, recall that any open subset Ω of $\mathbb{R}^n$ may be expressed as the union of a countable and increasing family of compact sets. That is, there is some family $\{K_i \subset \Omega \text{ compact} : i \in \mathbb{N}\}$ so that

$$\begin{array}{ll} \forall & i \in \mathbb{N} : \\ & K_i \subset K_{i+1} \end{array}$$

and

$$\Omega = \bigcup_{i \in \mathbb{N}} K_i$$

For each $i, j \in \mathbb{N}$ so that $i < j$ we may define the injective mapping

$$f_{i,j} : \mathcal{D}(K_i) \rightarrow \mathcal{D}(K_j)$$

through

$$f_{i,j}u : K_j \ni x \mapsto \begin{cases} u(x) & if \quad x \in K_i \\ 0 & if \quad x \notin K_i \end{cases}. \tag{1.21}$$

Furthermore, and in the same way as in (1.21), for each $i \in \mathbb{N}$ we may define the mapping

$$f_i : \mathcal{D}(K_i) \to \mathcal{D}(\Omega)$$

where

$$\mathcal{D}(\Omega) = \left\{ \varphi \in \mathcal{C}^\infty(\Omega) \middle| \begin{array}{ll} \exists & K \subset \Omega \text{ compact :} \\ & \varphi \in \mathcal{D}(K) \end{array} \right\}$$

so that the diagram

$$\begin{tikzcd} \mathcal{D}(K_i) \arrow[rr, "f_{i,j}"] \arrow[dr, "f_j"'] & & \mathcal{D}(K_j) \arrow[dl, "f_j"] \\ & \mathcal{D}(\Omega) & \end{tikzcd} \tag{1.22}$$

commutes whenever $i < j$. As such, and in view of the injectivity of the mappings $f_{i,j}$ and f_i, we may express $\mathcal{D}(\Omega)$ as the strict inductive limit of the inductive system $(\mathcal{D}(K_i), f_{i,j})_{i,j\in\mathbb{N}}$.

Each of the spaces $\mathcal{D}(K_i)$ carries in a natural way the structure of a Frechét space. Indeed, the family of seminorms $\{\rho_\alpha\}_{\alpha\in\mathbb{N}^n}$ defined through

$$\rho_\alpha : \mathcal{D}(K_i) \ni u \mapsto \sup\{|D^\alpha u(x)| : x \in K_i\},$$

where $\|\cdot\|_K$ denotes the uniform norm on $\mathcal{C}^0(K)$, defines a metrizable locally convex topology τ_i on $\mathcal{D}(K_i)$. As such, and in view of the construction of $\mathcal{D}(\Omega)$ as the strict inductive limit of the inductive system $(\mathcal{D}(K_i), f_{ij})$, the space $\mathcal{D}(\Omega)$ may be equipped with the locally convex strict inductive limit of the family of Fréchet spaces $(\mathcal{D}(K_i), \tau_i)_{i\in\mathbb{N}}$. An intuitive feeling for this topology on $\mathcal{D}(\Omega)$ may be obtained by considering convergent sequences. A sequence (φ_n) in $\mathcal{D}(\Omega)$ converges to $\varphi \in \mathcal{D}(\Omega)$ if and only if

$$\begin{array}{ll} \exists & K \subset\subset \Omega : \\ \forall & n \in \mathbb{N} : \\ & \mathrm{supp}\varphi_n \subseteq K \end{array}$$

and

$$\forall \quad \alpha \in \mathbb{N} : \\ \|D^{\alpha}\varphi - D^{\alpha}\varphi_n\|_K \to 0$$

The concept of weak derivative, such as in the sense of Sobolev [148], [149], is incorporated in the above functional analytic setting by associating with each $u \in L_2(\Omega)$ a continuous linear functional

$$T_u : \mathcal{D}(\Omega) \to \mathbb{R}$$

through

$$T_u : \varphi \mapsto \int_{\Omega} u(x)\varphi(x)\,dx. \tag{1.23}$$

As such, one obtains the inclusion

$$L_2(\Omega) \subseteq \mathcal{D}'(\Omega) = \left\{ T : \mathcal{D}(\Omega) \to \mathbb{R} \;\middle|\; \begin{array}{ll} 1) & T \text{ linear} \\ 2) & T \text{ continuous} \end{array} \right\}$$

In particular, if $u \in H^{2,m}(\Omega)$, then for each $|\alpha| \leq m$, we have

$$T_{D^{\alpha}u} : \varphi \mapsto \int_{\Omega} D^{\alpha}u(x)\varphi(x)\,dx = (-1)^{|\alpha|}\int_{\Omega} u(x)\,D^{\alpha}\varphi(x)\,dx. \tag{1.24}$$

Identifying with each $u \in H^{2,m}(\Omega)$ and every $|\alpha| \leq m$ the functional $D^{\alpha}T_u \in \mathcal{D}'(\Omega)$ which is defined as

$$D^{\alpha}T_u : \varphi \mapsto T_{D^{\alpha}u}\varphi \tag{1.25}$$

it is clear that $\mathcal{D}'(\Omega)$ also contains each weak partial derivative, up to order m, of functions in $H^{2,m}(\Omega)$. Generalizing the formula (1.24) to arbitrary continuous linear functionals in $\mathcal{D}'(\Omega)$, one may define generalized partial derivatives in $\mathcal{D}'(\Omega)$ of all orders for each $T \in \mathcal{D}'(\Omega)$ through

$$D^{\alpha}T : \mathcal{D}(\Omega) \ni \varphi \mapsto (-1)^{|\alpha|}T(D^{\alpha}\varphi) \in \mathbb{R} \tag{1.26}$$

Not each continuous linear functional on $\mathcal{D}(\Omega)$ can be described through (1.23). Indeed, suppose that $\Omega = \mathbb{R}^n$, and consider the linear functional δ, called the Dirac distribution, defined as

$$\delta : \mathcal{D}(\mathbb{R}^n) \ni \varphi \mapsto \varphi(0) \in \mathbb{R}. \tag{1.27}$$

Clearly (1.27) defines a continuous linear functional on $\mathcal{D}(\mathbb{R}^n)$. However, there is no locally integrable function u on $\mathbb{R}^n$ so that

$$\delta : \varphi \mapsto \int_{\Omega} u(x)\varphi(x)\,dx.$$

In this regard, for $a > 0$, consider the function $\varphi_a \in \mathcal{D}(\mathbb{R}^n)$ which is defined as

$$\varphi_a(x) = \begin{cases} e^{-\frac{1}{1-|x|/a}} & if \quad |x| < a \\ 0 & if \quad |x| \geq a \end{cases}$$

Suppose that there is an integrable function u on Ω that defines δ through (1.23). For each $a > 0$ we have

$$e^{-1} = \delta(\varphi_a) = \int_{\mathbb{R}^n} u(x)\varphi_a(x)\,dx \to 0$$

which is absurd. The space $\mathcal{D}'(\Omega)$ of continuous linear functionals on $\mathcal{D}(\Omega)$ is called the space of distributions on Ω and is denoted $\mathcal{D}'(\Omega)$. In view of the above example involving the Dirac distribution, it is clear that $\mathcal{D}'(\Omega)$ contains not only each of the Sobolev spaces $H^{2,m}(\Omega)$, for $m \geq 1$, but also more general generalized functions. Indeed, every locally integrable function u, that is, every Lebesgue measurable function that satisfies

$$\begin{array}{ll} \forall & K \subset \Omega \text{ compact}: \\ & \int_K |u(x)|dx < \infty \end{array}$$

may be associated with a suitable element T_u of $\mathcal{D}'(\Omega)$ in a canonical way.

The Schwartz linear theory of distributions has a rather natural position within the context of spaces of generalized functions which contain $\mathcal{C}^0(\Omega)$. In this regard, we should note that the main objective of the linear functional analytic approach of Schwartz [144] is the infinite differentiability of generalized functions, something that the sequential approach of Sobolev [148], [149] fails to achieve. In this regard, and in view of (1.26), it is clear that this aim is achieved within the setting of $\mathcal{D}'$ distributions. In fact, from the point of view of the existence of partial derivatives, the space $\mathcal{D}'(\Omega)$ posses a *canonical structure*. Indeed, in the chain of inclusions

$$\mathcal{C}^\infty(\Omega) \subset \ldots \subset \mathcal{C}^l(\Omega) \subset \ldots \subset \mathcal{C}^0(\Omega) \subset \mathcal{D}'(\Omega)$$

only $\mathcal{C}^\infty(\Omega)$ is closed under arbitrary partial derivatives in the classical sense. However, identifying $u \in \mathcal{C}^0(\Omega)$ with $T_u \in \mathcal{D}'(\Omega)$ through (1.23), we may again perform indefinite partial differentiation on u, with the partial derivatives defined as in (1.26), which, however, are no longer classical. Obviously, in view of (1.24) and (1.25), if $u \in \mathcal{C}^l(\Omega)$, then for $|\alpha| \leq l$, the partial derivative $D^\alpha u$ is the classical one, that is, the weak and classical derivatives coincide for sufficiently smooth functions.

The mentioned canonical structure of $\mathcal{D}'(\Omega)$ is the following:

$$\begin{array}{ll} \forall & T \in \mathcal{D}'(\mathbb{R}^n),\ K \subset \Omega \text{ compact}: \\ \exists & u \in \mathcal{C}^0(\Omega),\ \alpha \in \mathbb{N}^n: \\ & T_{|K} = D^\alpha u_{|K} \end{array} \tag{1.28}$$

Here D^α is the weak partial derivative (1.26). In other words, $\mathcal{D}'$ is a *minimal* extension of $\mathcal{C}^0$ in the sense that locally, every distribution is the weak partial derivative of a continuous function.

The linear theory of distributions, as shortly described above, as well as certain generalizations of it, see for instance [75], has proved to be a powerful tool in the study of PDEs, in particular in the case of linear, constant coefficient equations. Indeed, in view of the fact that $\mathcal{D}'(\Omega)$, as the dual of the locally convex space $\mathcal{D}(\Omega)$, is a vector space with the usual operations, and since $\mathcal{D}'(\Omega)$ contains $\mathcal{C}^\infty(\Omega)$ as a dense subspace, each constant coefficient linear partial differential operator

$$P(D) : \mathcal{C}^\infty(\Omega) \ni u \mapsto \sum_{|\alpha| \leq m} a_\alpha D^\alpha u \in \mathcal{C}^\infty(\Omega) \tag{1.29}$$

may be extended to the larger space $\mathcal{D}'(\Omega)$ through

$$P(D) : \mathcal{D}'(\Omega) \ni T \mapsto \sum_{|\alpha| \leq m} a_\alpha D^\alpha T \in \mathcal{D}'(\Omega)$$

In this regard, the first major result is due to Ehrenpreis [52] and Malgrange [103]. Namely, for each linear, constant coefficient partial differential operator (1.29), the generalized equation

$$P(D)T = \delta$$

admits a solution $T \in \mathcal{D}'(\Omega)$. From this it follows that, for any $\varphi \in \mathcal{D}(\Omega)$, the equation

$$P(D)u(x) = \varphi(x), \; x \in \Omega$$

with $P(D)$ defined as in (1.29), has a solution in $\mathcal{D}'(\Omega)$. This result alone justifies the use of the $\mathcal{D}'$-distributions in the study of linear, constant coefficient PDEs, and it has a variety of useful consequences and applications, see for instance [62] and [71].

In spite of the above mentioned canonical structure of the $\mathcal{D}'$-distributions in terms of partial differentiability, as well as the power of the linear theory of distributions in the context of linear, constant coefficient PDEs, the Schwartz distributions suffer from two major weaknesses. In the first place, we note that, for each $u \in \mathcal{C}^\infty(\Omega)$ and each $T \in \mathcal{D}'(\Omega)$, we can define the product of u and T in $\mathcal{D}'(\Omega)$ as

$$u \times T : \varphi \mapsto T(u \times \varphi) \tag{1.30}$$

That is, each distribution $T \in \mathcal{D}'(\Omega)$ can be multiplied with any $\mathcal{C}^\infty$-smooth function u. As such, and in view of the extension of the differential operators (1.26), every linear partial differential operator, say of order m, of the form

$$P(D)u(x) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha u(x)$$

where each coefficient a_α is $\mathcal{C}^\infty$-smooth, may be extended to the space $\mathcal{D}'(\Omega)$ of distributions on Ω. Indeed, in view of the linearity of the operator $P(D)$, we may, for any $T \in \mathcal{D}'(\Omega)$, define the distribution $P(D)T$ as

$$P(D)T : \varphi \mapsto \sum_{|\alpha| \leq m} a_\alpha D^\alpha T(\varphi)$$

As such, the partial differential equation

$$P(D)u(x) = g(x),$$

with $g \in \mathcal{C}^\infty(\Omega)$, defined by the operator $P(D)$, may be extended to a generalized equation in terms of distributions

$$P(D)T = T_g \tag{1.31}$$

where T_g is the distribution associated with the $\mathcal{C}^\infty$-smooth function g. Whenever the coefficient functions a_α are constant, that is,

$$\begin{array}{ll} \forall & |\alpha| \leq m : \\ \exists & a_\alpha \in \mathbb{R} : \\ & a_\alpha(x) = a_\alpha,\ x \in \Omega \end{array},$$

and the righthand term g has compact support, the generalized equation (1.31) admits a solution in $\mathcal{D}'(\Omega)$. Moreover, the existence of a solution holds also if the right hand term in (1.31) is any distribution with compact support. However, this result cannot be generalized to equations with nonconstant coefficients. In this regard, we may recall Lewy's impossibility result [97], see also [88]. Lewy showed that for a large class of functions $f_1, f_2 \in \mathcal{C}^\infty(\mathbb{R}^3)$, the system of first order linear PDEs

$$\begin{array}{rcl} -\frac{\partial}{\partial x_1}U_1 + \frac{\partial}{\partial x_2}U_2 - 2x_1\frac{\partial}{\partial x_3}U_2 - 2x_2\frac{\partial}{\partial x_3}U_1 & = & f_1 \\ -\frac{\partial}{\partial x_1}U_2 + \frac{\partial}{\partial x_2}U_1 + 2x_1\frac{\partial}{\partial x_3}U_1 - 2x_2\frac{\partial}{\partial x_3}U_2 & = & f_2 \end{array}, \tag{1.32}$$

which may be written as a single equation with complex coefficients, has no distributional solutions in any neighborhood of any point of $\mathbb{R}^3$. The most interesting aspect of the example (1.32) is that it is not the typical, esoteric, counterexample type of equation, but appears rather naturally in connection with the theory of functions of several complex variables, see for instance [88]. In view of the rather natural, not to mention simple, form of Lewy's example, and as will be elaborated upon further in the sequel, the Schwartz distributions prove to be insufficient from this point of view for large classes of equations.

Over and above the mere insufficiency of the Schwartz distributions from the point of view of the existence of generalized solutions to systems of PDEs, the space $\mathcal{D}'(\Omega)$ suffers from serious structural deficiencies. In particular, and since the early 1950s, it is known that $\mathcal{D}'(\Omega)$ does not admit any reasonable concept of multiplication that extends the usual pointwise multiplication of smooth functions. Indeed, Schwartz [145], see also [140], proved the following result.

Let $\mathcal{A}$ be an associative algebra so that $\mathcal{C}^0(\mathbb{R}) \subset \mathcal{A}$, and uv is the usual product of functions for each $u, v \in \mathcal{C}^0(\mathbb{R})$. If $D : \mathcal{A} \to \mathcal{A}$ is a differentiation operator, that is, D is linear and satisfies the Leibnitz rule for product derivatives, so that D restricted to $\mathcal{C}^1(\Omega) \subset \mathcal{A}$ is the usual differentiation operation, then there is no $\delta \in \mathcal{A}$, $\delta \neq 0$, so that

$$x\delta = 0. \tag{1.33}$$

This result is usually interpreted as follows. If $\delta \in \mathcal{D}'(\mathbb{R})$ is the Dirac delta distribution, then, in view of (1.30), for any $u \in \mathcal{C}^\infty(\mathbb{R})$ we have

$$u\delta : \mathcal{D}(\mathbb{R}) \ni \varphi \mapsto \delta(u\varphi) = u(0)\varphi(x) \tag{1.34}$$

Therefore, if we let $x \in \mathcal{C}^\infty(\mathbb{R})$ denote the identity function on $\mathbb{R}$, then from (1.34) it follows that

$$x\delta : \mathcal{D}(\mathbb{R}) \ni \varphi \mapsto \delta(x\varphi) = 0.$$

That is, $x\delta$ is the additive identity in $\mathcal{D}'(\mathbb{R})$. Therefore, in view of (1.33), it follows that $\delta = 0$. But $\delta \neq 0 \in \mathcal{D}'(\mathbb{R})$, and hence *there cannot be a suitable multiplication on* $\mathcal{D}'(\mathbb{R})$. In particular, this has the effect that one cannot formulate the concept of solution to nonlinear PDEs, such as (1.4) for instance, in the framework of the $\mathcal{D}'$ distributions.

In view of the above impossibility, in order to multiply arbitrary distributions in a *consistent* and *meaningful* way, it is necessary to find an embedding

$$\mathcal{D}'(\mathbb{R}) \hookrightarrow \mathcal{A} \tag{1.35}$$

where $\mathcal{A}$ is a suitable algebra wherein multiplication may be performed. However, in this regard, there typically occur misinterpretations of the mentioned Schwartz impossibility. Indeed, Schwartz's result is usually interpreted as stating that there cannot exists convenient algebras $\mathcal{A}$ such as in (1.35), and that a convenient multiplication of arbitrary distributions is not possible [71].

Furthermore, in view of the perceived impossibility of multiplying distributions in a convenient and meaningful way, it is widely believed that there can not be a *general* and *convenient* nonlinear theory of generalized functions. In particular, one cannot hope to obtain any significantly general and type independent theory for generalized solutions of nonlinear PDEs. It will be shown over and again in the sequel that this is in fact a misunderstanding, and suitable nonlinear theories of generalized functions can easily be constructed, see for instance [135] through [142], theories which deliver general existence and regularity results for large classes of systems of nonlinear PDEs.

Over and above the mentioned deficiencies of the Schwartz linear theory of distributions, both from the point of view of existence of generalized solutions to PDEs, as well as in terms of its capacity to handle also nonlinear problems, the space $\mathcal{D}'(\Omega)$ suffers from several other serious weaknesses. In this regard, we mention only that $\mathcal{D}'(\Omega)$ fails to be a flabby sheaf [140], a property that is fundamental in connection with the study of singularities, and that the use of the $\mathcal{D}'$ distributions is not convenient, from the point of view of exactness, to deal with sequential solutions, even to linear PDEs [140].

And now, taking into account the various weaknesses and deficiencies of the $\mathcal{D}'$ distributions, when it comes to the study of generalized solutions to both linear and nonlinear PDEs, there appears to be only two possible ways forward. In the first place, and as is commonly believed, it may seem that there cannot be a general and convenient framework for such generalized solutions, and instead various ad hoc methods may be applied to different equations. On the other hand, a suitable extension of the theory of distributions may be pursued, such as may be provided by suitable embeddings of $\mathcal{D}'(\Omega)$ into convenient algebras of generalized functions, such as in (1.35). These two alternatives have been pursued for the last five decades, as will be explained in the subsequent sections.

However, there is a third, if largely overlooked, possibility in pursuing a systematic account of generalized solutions to linear and nonlinear PDEs. In view of the above mentioned difficulties presented by the Schwartz distributions, in particular in connection with nonlinear PDEs, why should any theory of generalized solutions to systems of PDEs be restricted by the requirement that it must contain the $\mathcal{D}'$ distributions as a particular case? Indeed, one may start from the very beginning and ask what exactly are the requirements of such a theory? In this regard, there have lately been two independent attempts at such a completely *new* theory of generalized solutions to systems of PDEs, namely, the Order Completion Method [119], and the Central Theory for PDEs [115] through [118]. Both these theories, although approaching the subject of generalized solutions to PDEs from rather different points of view, have delivered general and type independent existence and regularity results for generalized solutions of large classes of systems of nonlinear PDEs.

1.2 Weak Solution Methods

In view of the deficiencies of the linear theory of $\mathcal{D}'$-distributions, in particular the impossibility of defining nonlinear operations on $\mathcal{D}'(\Omega)$, and the insufficiency of the $\mathcal{D}'$ framework of generalized functions in the context of the existence of generalized solutions of PDEs, it is widely held that there cannot be a convenient nonlinear theory of generalized functions and, moreover, a general and type independent theory for the existence and regularity of generalized solutions of systems of nonlinear PDEs is impossible [12].

In view of the above remarks concerning the failures of the linear theory of distributions, rather than pursuing the essential features that are at work in regards to the existence of generalized solutions to linear and nonlinear PDEs, perhaps within a context other than the usual functional analytic one, the typical approach to the problems of existence and regularity of generalized solutions of nonlinear PDEs consists of a collection of rather ad hoc methods, each developed with a particular equation, or at best a particular type of equation, in mind.

At this point it is worth noting that, ever since Sobolev [148], [149], the main, and to a large extent even exclusive, approach to solving linear and nonlinear PDEs has been that of functional analysis. In this regard, and as mentioned above, this approach consists of a collection of ad hoc methods, each applying to but a rather small class of equations. Furthermore, in particular in the case of nonlinear equations, this often leads to ill founded concepts of a solution of such equations.

In this section we will briefly discuss the general framework of the more popular such methods for solving linear and nonlinear systems of equations, namely, the mentioned weak solution methods, see for instance [54] or [99], [100]. Furthermore, those particular difficulties that arise when applying such methods to nonlinear equations will be indicated [140].

In this regard, let us, at first, consider a linear partial differential equation of order m

$$L\left(D\right)u\left(x\right)=f\left(x\right),\ x\in\Omega \tag{1.36}$$

where $f:\Omega\to\mathbb{R}$ is, say, a continuous function, and the partial differential operator is of the form

$$L\left(D\right)u\left(x\right)=\sum_{|\alpha|\leq m}a_{\alpha}\left(x\right)D^{\alpha}u\left(x\right), \tag{1.37}$$

where $a_{\alpha}:\Omega\to\mathbb{R}$ are sufficiently smooth functions, for instance $a_{\alpha}\in\mathcal{C}^{0}\left(\Omega\right)$. With the partial differential operator $L\left(D\right)$ one associates a mapping

$$L:\mathcal{A}\to\mathcal{B} \tag{1.38}$$

where $\mathcal{A}$ is a vector space of sufficiently smooth functions on Ω, such as $\mathcal{A}=\mathcal{C}^{m}\left(\Omega\right)$, and $\mathcal{B}$ is an appropriate linear space of functions which contains the righthand term f.

The essential idea behind the so called weak solution methods is to construct an infinite sequence of partial differential operators $L_{i}\left(D\right)$ which approximate (1.37) in a suitable sense, so that each of the infinite sequence of PDEs

$$L_{i}\left(D\right)u_{i}\left(x\right)=f\left(x\right),\ x\in\Omega \tag{1.39}$$

admits a classical solution

$$u_{i}\in\mathcal{A},\ i\in\mathbb{N} \tag{1.40}$$

With a suitable choice for the approximate partial differential operators $\{L_i : i \in \mathbb{N}\}$ in (1.39), and solutions $\{u_i : i \in \mathbb{N}\}$ to (1.39), as well as appropriate linear space topologies on $\mathcal{A}$ and $\mathcal{B}$, one has

$$(Lu_i) \text{ converges to } f \text{ in } \mathcal{B}. \tag{1.41}$$

Furthermore, using some compactness, monotonicity or fixed point argument, one may often extract a Cauchy sequence from the sequence $(u_i) \subset \mathcal{A}$, which we denote by (u_i) as well. Thus, we have now obtained a sequence (u_i) in $\mathcal{A}$ so that

$$(u_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp \tag{1.42}$$

and

$$(Lu_i) \text{ converges to } f \in \mathcal{B}, \tag{1.43}$$

where $\mathcal{A}^\sharp$ denotes the completion of $\mathcal{A}$ in its given vector space topology. Now, in view of (1.42) and (1.43), $u^\sharp \in \mathcal{A}^\sharp$ is considered a generalized solution to (1.36).

Note that, in view of the typical difficulties involved in the steps (1.39) to (1.43), in particular when initial or boundary value problems are associated with the PDE (1.36), one ends up with very few Cauchy sequences in (1.42), if in fact not with a *single* such sequence. Furthermore, based solely on the very few, if in fact not the single Cauchy sequence in (1.42) and (1.43), the partial differential operator (1.37) is extended to a mapping

$$L(D) : \{u^\sharp\} \cup \mathcal{A} \rightarrow \mathcal{B}. \tag{1.44}$$

As such, this customary method for finding generalized solutions to PDEs amounts to nothing but an ad hoc, pointwise extension of the partial differential operator $L(D)$.

The deficiency of the above solution method is clear. Indeed, the extension (1.44) of the partial differential operator $L(D)$ is based on only very few Cauchy sequence (1.42) in $\mathcal{A}$. Moreover, the sequence in (1.42), and therefore also the generalized solution $u^\sharp$ of (1.36), is often obtained by an arbitrary subsequence selection from (1.40).

In the case of linear PDEs such as (1.36), the rather objectionable construction of a generalized solution to such an equation may, to some extend, be justified. Indeed, in this case, with the coefficients a_α in (1.37) sufficiently smooth, it happens that due to the phenomenon of *automatic continuity* of certain classes of linear operators, the above construction of a generalized solution to (1.36) is valid. Indeed, suppose that we obtain a Cauchy sequence (u_i) in $\mathcal{A}$ so that

$$(u_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp \tag{1.45}$$

and

$$(L(D)\, u_i) \text{ converges to } f \in \mathcal{B}. \tag{1.46}$$

Given now any sequence (v_i) in $\mathcal{A}$ so that

$$(v_i) \text{ converges to } 0 \in \mathcal{A}$$

we may define the Cauchy sequence in $\mathcal{A}$

$$(w_i) = (u_i + v_i) .$$

From the linearity of the operator $L(D)$ we now obtain

$$(L(D)\, w_i) = (L(D)\, u_i) + (L(D)\, v_i) \tag{1.47}$$

If, as most often happens to be the case, the mapping $L(D)$ is continuous, we have

$$(L(D)\, v_i) \text{ converges to } 0 \in \mathcal{B}. \tag{1.48}$$

Now, in view of (1.47) and (1.48), it follows that

$$(L(D)\, w_i) \text{ converges to } f \in \mathcal{B}$$

Therefore, based solely on the single Cauchy sequence in (1.45) and (1.46), we have

$$\begin{array}{ll} \forall & (u_i) \subset \mathcal{A} : \\ & (u_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp \Rightarrow (L(D)\, u_i) \text{ converges to } f \end{array} . \tag{1.49}$$

The relationships (1.47) to (1.49) affirms the interpretation of $u^\sharp \in \mathcal{A}^\sharp$ as a generalized solution to (1.36).

In contradistinction with the case of linear PDEs, when the procedure (1.39) through (1.43) for establishing the existence of weak solutions is applied to a nonlinear PDE

$$T(D)\, u(x) = f(x)\, ,\; x \in \Omega \tag{1.50}$$

one critical point is often overlooked, namely, the nonlinear operator $T(D)$ is typically not compatible with the vector space topologies on $\mathcal{A}$ and B. In this case the claim that $u^\sharp \in \mathcal{A}^\sharp$ is a generalized solution to (1.50) is rather objectionable. Indeed, in the case of a nonlinear PDE (1.50) there is in fact a double breakdown in (1.45) to (1.49). In this regard, note that in case the linear partial differential operator $L(D)$ in (1.36) is replaced with a nonlinear operator, such as in (1.50), both the crucial steps in (1.47) and (1.48) will in general break down. In that case, then, we cannot in general deduce from very few, if not in fact one single Cauchy sequence $(u_i) \subset \mathcal{A}$ such that

$$(u_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp \tag{1.51}$$

and

$$(T(D)\, u_i) \text{ converges to } f \in \mathcal{B}, \tag{1.52}$$

that

$$\begin{array}{ll} \forall & (u_i) \subset \mathcal{A} : \\ & (u_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp \Rightarrow (L(D)\,u_i) \text{ converges to } f \end{array} .$$

In view of this double breakdown in the customary weak solution methods, when applied to nonlinear problems, it is clear that such methods are typically ill founded when applied to such nonlinear problems.

It is exactly this double breakdown which, for a long time, was usually overlooked when applying solution methods for linear PDEs to nonlinear ones. In part, this oversight is perhaps due to the fact that, in the particular case of linear PDEs, the method (1.39) to (1.43) happens to be correct. However, and in view of the above remarks, it is clear that the extension of linear methods to nonlinear problems often require essentially new ways of thinking. For an excellent survey of the difficulties of several such well known extensions can be found in [140] or [165].

The careless application of essentially linear methods such as in (1.39) to (1.43) to nonlinear problems can, and in fact often does, lead to absurd conclusions, see for instance [140]. In this regard, a most simple example is given by the zero order nonlinear system of equations

$$\begin{array}{rcl} u & = & 0 \\ & & \\ u^2 & = & 1 \end{array} \tag{1.53}$$

which, using the customary weak solution method (1.39) to (1.43), admit both weak and strong solutions, so that, apparently, we have proven the blatant absurdity

$$0 = 1 \text{ in } \mathbb{R}.$$

Indeed, if we set $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R})$ with the topology induced by $\mathcal{D}'(\mathbb{R})$, and we take $\mathcal{B} = \mathcal{D}'(\mathbb{R})$, then the sequence (v_i) defined as

$$v_i(x) = \sqrt{2}\cos(ix)\,,\ i \in \mathbb{N}$$

is Cauchy in $\mathcal{A}$ and

$$(v_i) \text{ converges to } u = 0$$

both weakly and strongly in $\mathcal{A}^\sharp = \mathcal{D}'(\mathbb{R})$. Furthermore, we shall also have

$$\left(v_i^2\right) \text{ converges to } v = 1$$

both *weakly and strongly* in $\mathcal{B} = \mathcal{D}'(\mathbb{R})$. That is, the sequences (v_i) and (v_i^2) converge to their respective limits with respect to the usual topology on $\mathcal{D}'$ and

$$\begin{array}{lll} \forall & \varphi \in \mathcal{D}(\mathbb{R}) : & \\ & 1) & T_{v_i}(\varphi) \to 0 \text{ in } \mathbb{R} \\ & 2) & T_{v_i^2}(\varphi) \to 1 \text{ in } \mathbb{R} \end{array} . \tag{1.54}$$

As such, according to the typical weak solution method (1.39) to (1.43), the sequence (v_i) defines both a weak and strong solution to (1.53).

Let us now take a closer look at how the above nonlinear stability paradox, namely, the existence of both weak and strong solutions to (1.53), effects the customary sequential approach (1.39) to (1.43). In this regard, consider the nonlinear partial differential operator given by

$$T(D)\,u = L(D)\,u + u^2$$

where $L(D)$ is a linear partial differential operator such as in (1.37). Suppose that the system (1.53) has a solution in $\mathcal{A}$, that is,

$$\begin{array}{l} \exists \quad (v_i) \subset \mathcal{A} \text{ a Cauchy sequence :} \\ \quad 1) \quad (v_i) \text{ converges to } 0 \text{ in } \mathcal{A} \\ \quad 2) \quad (v_i^2) \text{ converges to } 1 \text{ in } \mathcal{B} \end{array} . \tag{1.55}$$

Let (u_i) be the Cauchy sequence in (1.51) and (1.52). Define the Cauchy sequence (w_i) in $\mathcal{A}$ as

$$w_i = u_i + \lambda v_i,\ i \in \mathbb{N},$$

for an arbitrary but fixed $\lambda \in \mathbb{R}$. In view of (1.51) and (1.55) we have

$$(w_i) \text{ converges to } u^\sharp \in \mathcal{A}^\sharp. \tag{1.56}$$

Hence the sequence (w_i) defines the same generalized function as (u_i) in (1.51). Now in view of (1.52) it follows that

$$T(D)\,w_i = T(D)\,u_i + \lambda L(D)\,v_i + 2\lambda u_i v_i + \lambda^2 v_i^2.$$

Hence, assuming that $(u_i v_i)$ is a Cauchy sequence in $\mathcal{A}$ with limit $v^\sharp \in \mathcal{A}^\sharp$, (1.52) and (1.55) yield

$$(T(D)\,w_i) \text{ converges to } f + \lambda^2 + 2\lambda v^\sharp \text{ in } \mathcal{B} \tag{1.57}$$

if, as it often happens, $(L(D)\,v_i)$ converges to 0 in $\mathcal{B}$, such as for instance in the topology of $\mathcal{D}'(\Omega)$ when $L(D)$ has $\mathcal{C}^\infty$-smooth coefficients.

Since $\lambda \in \mathbb{R}$ is arbitrary, (1.56) and (1.57) yields

$$T(D)\,u^\sharp \neq f \tag{1.58}$$

In this way, the same $u^\sharp \in \mathcal{A}^\sharp$ that is a generalized solution to the equation $T(D)\,u = f$ according to the customary interpretation of (1.51) and (1.52), now, in view of (1.58), no longer happens to be a solution to this equation.

In view of the above example, it is clear that, as far as *nonlinear* partial differential equations are concerned, the extension of the concept of classical solution to that of the concept of generalized solution along the lines (1.39) to (1.43), is an *improper* generalization of various classical extensions, such as for instance the extension of the rational numbers into the real numbers.

Over the last thirty years or so, there has been a limited awareness among the functional analytic school concerning the difficulties involved in extending linear methods to nonlinear problems, see for instance [14], [44], [48], [127], [147] and [151]. However, the techniques developed to overcome these difficulties, such as the Tartar-Murat compensated compactness and the Young measure associated with weakly convergent sequences of functions subject to differential constraints on an algebraic manifold, can deal only with particular types of nonlinear PDEs and sequential solutions. The effect of this limited approach is an obfuscation of the basic underlying reasons, whether these are of an algebraic or topological nature, for such problems as the nonlinear stability paradox discussed in this section. Not to mention that there is no attempt to develop a systematic nonlinear theory of generalized functions that would be able to accommodate large classes of nonlinear PDEs.

1.3 Differential Algebras of Generalized Functions

As we have mentioned at the end of Section 1.1, and as demonstrated in Section 1.2, in order to solve large classes of linear and nonlinear PDEs, it is necessary to go beyond the usual functional analytic methods for PDEs, including the Schwartz $\mathcal{D}'$ distributions. In particular, the Schwartz impossibility result places serious restrictions on use of distributions in the study of nonlinear PDEs. As such, in order to obtain a theory of generalized functions that would be able to handle large classes of nonlinear PDEs, and at the same time contain the $\mathcal{D}'$ distributions, the underlying reasons for the mentioned Schwartz impossibility result, reasons that turn out to be rather algebraic than topological [140], should be clearly understood.

In this regard [140], the difficulties involved in establishing a suitable framework of generalized functions for the solutions of nonlinear partial differential equations may be viewed as a consequence of a basic algebraic conflict between the trio of discontinuity, multiplication and differentiation. In order to illustrate how such a conflict might arise, we consider the most simple discontinuous function, namely, the Heaviside function $H : \mathbb{R} \rightarrow \mathbb{R}$ given by

$$H(x) = \begin{cases} 0 & if \quad x \leq 0 \\ \\ 1 & if \quad x > 0 \end{cases} \tag{1.59}$$

When a discontinuous function such as H in (1.59) appears as a solution to a nonlinear PDEs, this function will necessarily be subjected to the operations of differentiation and multiplication. As such, a natural and intuitive setting for a theory of generalized functions that would include, in particular, the discontinuous function in (1.59), would be a ring of functions

$$\mathcal{A} \subseteq \{u : \mathbb{R} \rightarrow \mathbb{R}\} \tag{1.60}$$

so that

$$H \in \mathcal{A}. \tag{1.61}$$

Furthermore, there should be a differential operator

$$D : \mathcal{A} \to \mathcal{A} \tag{1.62}$$

defined on $\mathcal{A}$. That is, D is a linear operator that satisfies the Leibniz rule for product derivatives

$$D(uv) = uDv + vDu \tag{1.63}$$

Already within this basic setup (1.60) through (1.63) we encounter rather surprising and undesirable consequences.

Indeed, in view of (1.59) and (1.60), it follows that

$$\begin{array}{ll} \forall & m \in \mathbb{N},\, m \geq 1 : \\ & H^m = H \end{array} \tag{1.64}$$

Furthermore, it follows from (1.60) that $\mathcal{A}$ is both associative and commutative. Hence (1.63) and (1.64) implies the relation

$$\begin{array}{ll} \forall & m \in \mathbb{N},\, m \geq 2 : \\ & mH \cdot DH = DH \end{array} . \tag{1.65}$$

From (1.65) it now follows that

$$\begin{array}{ll} \forall & p,\, q \in \mathbb{N},\, p, q \geq 2 : \\ & p \neq q \Rightarrow \left(\frac{1}{p} - \frac{1}{q}\right) DH = 0 \in \mathcal{A} \end{array}$$

which implies

$$DH = 0 \in \mathcal{A}. \tag{1.66}$$

However, in view of the fact that our theory should contain the $\mathcal{D}'$ distributions we have that

$$\mathcal{D}'(\mathbb{R}) \subset \mathcal{A} \tag{1.67}$$

and the differential operator D on $\mathcal{A}$ extends the distributional derivative. In particular, it follows from (1.67) that

$$\delta \in \mathcal{A}$$

where δ is the Dirac distribution. Furthermore, since $H \in \mathcal{D}'(\mathbb{R})$ and

$$DH = \delta \in \mathcal{D}'(\mathbb{R}),$$

it follows by (1.66) that

$$\delta = 0 \in \mathcal{A} \tag{1.68}$$

which is of course false.

It is now obvious that if we wish to define nonlinear operations, in particular unrestricted multiplication, on generalized functions that contain the $\mathcal{D}'$ distributions in a consistent and useful way, some of the assumptions (1.60) to (1.63) must be relaxed. This can be done in any of several different ways.

Indeed, while the algebra $\mathcal{A}$ should contain functions such as the Heaviside function H in (1.59), it need not be an algebra of functions from $\mathbb{R}$ to $\mathbb{R}$. That is, $\mathcal{A}$ may contain elements that are more general than such functions $u : \mathbb{R} \rightarrow \mathbb{R}$. Furthermore, the multiplication in $\mathcal{A}$ need not be so closely related to multiplication of real valued functions. As such, (1.60) need not necessarily hold.

Regarding the differentiation operator D in (1.63), it is important to note that the requirement (1.62) is of highly restrictive assumption. Indeed, (1.62) implies that each $u \in \mathcal{A}$ is indefinitely differentiable. That is,

$$\begin{array}{ll} \forall & m \in \mathbb{N}, m \geq 1 : \\ \forall & u \in \mathcal{A} : \\ & D^m u \in \mathcal{A} \end{array} .$$

Of course, this is the case if $\mathcal{A} \subseteq \mathcal{C}^\infty(\mathbb{R})$, which, in view of (1.61), is not possible. As such, one may also want to keep in mind the possibility that the differential operator may rather be defined as

$$D : \mathcal{A} \rightarrow \overline{\mathcal{A}}$$

where $\overline{\mathcal{A}}$ is *another* algebra of generalized functions. In this case, the Leibniz Rule (1.63) may be preserved in this more general situation by assuming the existence of an *algebra homomorphism*

$$\mathcal{A} \ni u \mapsto \overline{u} \in \overline{\mathcal{A}}$$

and rewriting (1.63) as

$$D(u \cdot v) = (Du) \cdot \overline{v} + \overline{u} \cdot (Dv) \tag{1.69}$$

where the product on the left of (1.69) is taken in $\mathcal{A}$, and the product on the right is taken in $\overline{\mathcal{A}}$.

The above two relaxations, namely, on the algebra $\mathcal{A}$ and the derivative operator D are sufficient in order to obtain generalized functions which extend the $\mathcal{D}'$ distributions and admit generalized solutions to large classes of linear and nonlinear PDEs. Furthermore, the arguments leading to (1.66) are purely of an algebraic nature, and do not involve calculus or topology. This is precisely the reason for the power and usefulness of the so called 'algebra first' approach, [140]. We will shortly describe such an approach to constructing generalized functions, initiated by Rosinger [135], [136] and developed further in [137], [138] and [140].

In this regard, it is helpful to recall that generalized solutions to linear PDEs are typically constructed as elements of the completion of a suitably chosen locally convex topological, in particular metrizable, vector space $\mathcal{A}$ of sufficiently smooth functions $u : \Omega \rightarrow \mathbb{R}$, see Section 1.2. From a more abstract point of view, the completion $\mathcal{A}^\sharp$ of the metrizable topological vector space $\mathcal{A}$ may be constructed as

$$\mathcal{A}^\sharp = \mathcal{S}/\mathcal{V} \tag{1.70}$$

where

$$\mathcal{V} \subset \mathcal{S} \subset \mathcal{A}^{\mathbb{N}} \tag{1.71}$$

with $\mathcal{A}^{\mathbb{N}}$ the set of sequences in $\mathcal{A}$ and $\mathcal{S}$ is the set of all Cauchy sequences in $\mathcal{A}$. With termwise operations on sequences, $\mathcal{A}^{\mathbb{N}}$ is in a natural way a vector space, while $\mathcal{S}$ and $\mathcal{V}$ are suitable vector subspaces of $\mathcal{A}^{\mathbb{N}}$. Therefore, the quotient space $\mathcal{S}/\mathcal{V}$ is again a vector space.

In the case of *nonlinear* PDEs, we shall instead be interested in a suitable algebra of generalized functions. In this regard, the above construction in (1.70) to (1.71) may be adapted for that purpose. Indeed, we may choose a suitable subalgebra of smooth functions

$$\mathcal{A} \subseteq \mathcal{C}^m(\Omega) \tag{1.72}$$

and

$$\mathcal{I} \subset \mathcal{S} \subset \mathcal{A}^{\mathbb{N}} \tag{1.73}$$

where $\mathcal{S}$ is a subalgebra in $\mathcal{A}^{\mathbb{N}}$, while $\mathcal{I}$ is an ideal in $\mathcal{S}$. Then the quotient algebra

$$\mathcal{A}^\sharp = \mathcal{S}/\mathcal{I} \tag{1.74}$$

can offer the representation for our algebra of generalized functions. Furthermore, when

$$\mathcal{A} \subseteq \mathcal{C}^\infty(\Omega),$$

then for suitable choices of the subalgebra $\mathcal{S} \subset \mathcal{A}^{\mathbb{N}}$ and the ideal $\mathcal{I} \subset \mathcal{S}$ in (1.72) to (1.74), there is a *vector space embedding*

$$\mathcal{D}'(\Omega) \rightarrow \mathcal{A}^\sharp \tag{1.75}$$

However, in general the embedding (1.75) cannot be achieved in a canonical way, as can be seen by straightforward ring theoretic arguments connected with the existence of maximal off diagonal ideals [140]. Colombeau [39], [40] constructed such an algebra of generalized functions that admits a canonical embedding (1.75).

Now it should be noted that, in view of (1.72), the algebra $\mathcal{A}$ is automatically both associative and commutative. Therefore, with the termwise operations on sequences, $\mathcal{A}^{\mathbb{N}}$ is also associative and commutative so that the algebra $\mathcal{A}^{\sharp}$ of generalized functions in (1.74) will also be associative and commutative.

At first glance it may appear that the construction in (1.72) to (1.74) is rather arbitrary. However, see for instance [140], such concerns may be addressed and clarified to a good extent. In particular, the following points may be noted.

First of all, it is a natural condition to impose on the differential algebra $\mathcal{A}$ that

$$\mathcal{A} \subset \mathcal{A}^{\sharp}. \tag{1.76}$$

In particular, this addresses the issue of *consistency* of usual classical solutions to a nonlinear PDEs with generalized solutions in $\mathcal{A}^{\sharp}$. In this regard, the purely algebraic *neutrix condition* will particularize the above framework (1.72) to (1.74) so as to incorporate (1.76). This purely algebraic condition characterizes the requirement (1.76), yet it has proven to be surprisingly powerful.

In this regard, recall that in the case of a general vector space $\mathcal{A}$ of usual functions $u : \Omega \to \mathbb{R}$ we have taken vector subspaces $\mathcal{V} \subset \mathcal{S} \subset \mathcal{A}^{\mathbb{N}}$ and have defined $\mathcal{A}^{\sharp} = \mathcal{S}/\mathcal{V}$ as a space whose elements $u^{\sharp} \in \mathcal{A}^{\sharp}$ generalize the classical functions $u \in \mathcal{A}$. As such, one should have a *vector space embedding*

$$\mathcal{A} \subset \mathcal{A}^{\sharp} = \mathcal{S}/\mathcal{V} \tag{1.77}$$

which is defined by a linear injection

$$\iota_{\mathcal{A}} : \mathcal{A} \ni u \mapsto \iota_{\mathcal{A}}(u) = (u) + \mathcal{V} \in \mathcal{A}^{\sharp} \tag{1.78}$$

where (u) is the constant sequence with all terms equal to $u \in \mathcal{A}$.

We may reformulate (1.77) to (1.78) in the following more convenient way. Let $\mathcal{O} \subset \mathcal{A}^{\mathbb{N}}$ denote the null vector subspace, that is, the subspace consisting only of the constant zero sequence, and let

$$\mathcal{U}_{\mathcal{A},\mathbb{N}} = \{\iota_{\mathcal{A}}(u) \,:\, u \in \mathcal{A}\} \tag{1.79}$$

be the vector subspace in $\mathcal{A}^{\mathbb{N}}$ consisting of all constant sequences in $\mathcal{A}$. That is, $\mathcal{U}_{\mathcal{A},\mathbb{N}}$ is the *diagonal* in the cartesian product $\mathcal{A}^{\mathbb{N}}$. Then (1.77) to (1.78) is equivalent to the commutative diagram

$$\begin{tikzcd}
\mathcal{V} \arrow[r, "\subset"] & \mathcal{S} \arrow[r, "\subset"] & \mathcal{A}^{\mathbb{N}} \\
\mathcal{O} \arrow[u, "\subset"] \arrow[r, "\subset"'] & \mathcal{U}_{\mathcal{A},\mathbb{N}} \arrow[u, "\subset"'] &
\end{tikzcd} \tag{1.80}$$

together with the co called *off diagonal* condition

$$\mathcal{V} \cap \mathcal{U}_{\mathcal{A},\mathbb{N}} = \mathcal{O} \tag{1.81}$$

which is called the *neutrix condition* [140], the name being suggested by similar ideas introduced in [153] within a so called 'neutrix calculus' developed in connection with asymptotic analysis. Following the terminology in [153], see also [140], a sequence $(u_i) \in \mathcal{V}$ is called $\mathcal{V}$-negligible. In this sense, for two functions $u, v \in \mathcal{A}$, their difference $u - v \in \mathcal{A}$ is $\mathcal{V}$-negligible if and only if $\iota_{\mathcal{A}}(u) - \iota_{\mathcal{A}}(v) = \iota_{\mathcal{A}}(u - v) \in \mathcal{V}$, which in view of (1.81) is equivalent to $u = v$. As such, the neutrix condition (1.81) simply means that the quotient structure $\mathcal{S}/\mathcal{V}$ distinguishes between classical functions in $\mathcal{A}$.

Now, as mentioned, in the case of nonlinear PDEs we may be interested in constructing algebras of generalized functions such as is done in (1.72) to (1.74). In particular, (1.79) to (1.81) may be reproduced in this setting, given an *algebra* $\mathcal{A}$ of sufficiently smooth functions $u : \Omega \to \mathbb{R}$, a subalgebra $\mathcal{S}$ of $\mathcal{A}^{\mathbb{N}}$, and an ideal $\mathcal{I} \subset \mathcal{S}$, such that the inclusion diagram

$$\begin{tikzcd}
\mathcal{I} \arrow[r, "\subset"] & \mathcal{S} \arrow[r, "\subset"] & \mathcal{A}^{\mathbb{N}} \\
\mathcal{O} \arrow[u, "\subset"] \arrow[r, "\subset"'] & \mathcal{U}_{\mathcal{A},\mathbb{N}} \arrow[u, "\subset"'] &
\end{tikzcd} \tag{1.82}$$

satisfies the off diagonal or neutrix condition

$$\mathcal{I} \cap \mathcal{U}_{\mathcal{A},\mathbb{N}} = \mathcal{O}. \tag{1.83}$$

It follows that, similar to (1.77) through (1.78), for every *quotient algebra* $\mathcal{A}^\sharp = \mathcal{S}/\mathcal{I}$ we have the *algebra embedding*

$$\iota_{\mathcal{A}} : \mathcal{A} \ni u \mapsto \iota_A(u) = (u) + \mathcal{I} \in \mathcal{A}^\sharp \tag{1.84}$$

Furthermore, the conditions (1.80) to (1.83) are again necessary and sufficient for (1.84).

The neutrix condition, although of a simple and purely algebraic nature, turns out to be highly important and powerful in the study of such algebras of generalized functions. In particular, variants of this condition characterize the existence and structure of so called *chains of differential algebras* [138], [140]. Furthermore, the neutrix condition determines the structure of ideals $\mathcal{I}$ which play an important role in the stability, exactness and generality properties of the algebras of generalized functions.

The algebras of generalized functions $\mathcal{A}^\sharp$ constructed in (1.72) to (1.74) may be further particularized by introducing the natural requirement that nonlinear partial differential operators should be extendable to such algebras. In this regard, let us consider a polynomial type nonlinear partial differential operator

$$T(D) : \mathcal{C}^\infty(\Omega) \ni u \mapsto \sum_{i=1}^{h} c_i \prod_{j=1}^{k_i} D^{p_{ij}} u \in \mathcal{C}^\infty(\Omega) \tag{1.85}$$

where for $1 \leq i \leq h$ we have $c_i \in \mathcal{C}^\infty(\Omega)$. For convenience, we shall consider the problem of extending the partial differential operator (1.85) to a differential algebra

$$P(D) : \mathcal{A}^\sharp \rightarrow \mathcal{A}^\sharp \tag{1.86}$$

where the original algebra of classical functions satisfies

$$\mathcal{A} \subseteq \mathcal{C}^\infty(\Omega). \tag{1.87}$$

In this regard, we note that in order to obtain an extension (1.86), it is sufficient to extend the usual partial differential operators to mappings

$$D^p : \mathcal{A}^\sharp \rightarrow \mathcal{A}^\sharp,\ p \in \mathbb{N}^n. \tag{1.88}$$

This can easily be done by making the assumption

$$\begin{array}{l} \forall \quad p \in \mathbb{N}^n : \\ \qquad 1) \quad D^p\mathcal{I} \subset \mathcal{I} \\ \qquad 2) \quad D^p\mathcal{S} \subset \mathcal{A} \end{array} \tag{1.89}$$

In this case (1.86) can be defined as

$$D^p U = D^p s + \mathcal{I} \in \mathcal{A}^\sharp = \mathcal{A}/\mathcal{I},\ p \in \mathbb{N}^n \tag{1.90}$$

for each

$$U = s + \mathcal{I} \in \mathcal{A}^\sharp = \mathcal{S}/\mathcal{I}$$

where for every sequence $s = (u_n) \in \mathcal{A}^{\mathbb{N}}$ we define

$$D^p s = (D^p u_n)\,,\ p \in \mathbb{N}^n. \tag{1.91}$$

The extension (1.86) of the nonlinear partial differential operator (1.85) can now be obtained as follows. For a given

$$U = s + \mathcal{I} \in \mathcal{A}^\sharp = \mathcal{S}/\mathcal{I},\ s = (u_n) \in \mathcal{S}$$

we define

$$T(D)\,U = T(D)\,t + \mathcal{I} \in \mathcal{S}/\mathcal{I} = \mathcal{A}^\sharp$$

where

$$t = (v_n) \in \mathcal{S},\ t - s \in \mathcal{I}. \tag{1.92}$$

The construction (1.88) to (1.92) of the extension (1.88) may be replicated within a far more general setting. Indeed, the nonlinear operations, such as multiplication, are performed only in the range of the partial differential operator (1.85). As such, only the range need be an algebra, while we are more free in choosing the domain. In particular, we can replace the extensions (1.88) with

$$D^p : \mathcal{E}^\sharp \to \mathcal{A}^\sharp,\ p \in \mathbb{N}^n$$

where $\mathcal{E}^\sharp$ is a suitable *vector space* of generalized functions constructed in (1.77) to (1.78).

Furthermore, since the partial differential operator (1.85) is of finite order, say m, one may relax the condition (1.87) by only requiring

$$\mathcal{A} \subseteq \mathcal{C}^m(\Omega)$$

while the assumption (1.89) must now obviously be replaced with

$$\begin{array}{ll} \forall & p \in \mathbb{N}^n,\ |p| \le m: \\ & 1)\quad D^p\mathcal{I} \subset \mathcal{I} \\ & 2)\quad D^p\mathcal{S} \subset \mathcal{A} \end{array} . \tag{1.93}$$

The algebras of generalized functions can be further particularized in connection with the important concepts of *generality*, *exactness* and *stability* [140].

The method described in this section for constructing algebras of generalized functions turns out to be particularly efficient in the study of nonlinear PDEs. In this regard, the particular version of this theory, developed in [39], [40] has proved to be highly successful, both in connection with the exact and numerical solutions to nonlinear PDEs. However, the more general version of the theory developed by Rosinger [135], [136], [137], [138] and [140] provides a better insight into the structure of what may be conveniently called *all possible algebras of generalized functions*. Furthermore, the more general theory in [140] has delivered deep results such as a global version of the Cauchy-Kovalevskaia Theorem [141]. Such a result cannot be replicated within the more specific framework of the Colombeau Algebras. This is due to the polynomial type growth conditions imposed on the generalized functions in Colombeau's algebras. In particular, in this case the subalgebra in (1.73) is not equal to $\mathcal{A}^{\mathbb{N}}$. As such, one cannot define arbitrary smooth operations on generalized functions, since an arbitrary analytic function may grow faster than any polynomial.

1.4 The Order Completion Method

We have already mentioned the well known but often overlooked, if in fact not ignored, fact that the Schwartz distributions, and other similar spaces of generalized functions, suffer from certain *structural* weaknesses, and that these spaces fail to contain generalized solutions to a significantly general class of systems of PDEs. In particular, one may recall the Schwartz Impossibility Result, the Lewy Impossibility Result, as well as the nonlinear stability paradox discussed in Section 1.2. In view of such weaknesses, and as mentioned earlier, in order to develop a *general* framework for generalized solutions to PDEs, it is crucial that one goes beyond the usual distributions and related linear spaces that are customary in the study of nonlinear PDEs.

In this regard, one may then chose to construct spaces in such a way as to extend the Schwartz distributions. On the other hand, in view of the insufficiency and structural weaknesses of the distributions, one may start all over, and construct convenient spaces of generalized solutions which do not necessarily contain the distributions. The approach mentioned first is pursued in its full generality through Rosinger's Algebra First approach, as developed in [135] through [142], and discussed in Section 1.3, as well as the particular, yet highly important, case of that theory developed by Colombeau [39], [40]. The second possibility, that is, to define spaces of generalized functions without reference to the Schwartz distributions and other customary linear spaces of generalized functions, was pursued in a *systematic* and *general* way for the first time in [119], where spaces of generalized solutions are constructed through the process of Dedekind Order Completion of spaces of usual smooth functions on Euclidean domains.

The advantage of this approach, in comparison with the usual functional analytic methods, in particular as far as its generality and type independent power is concerned, comes from the fact that it is formulated within the context of the more basic concept of order. In this regard, we may note that present day mathematics is a multi layered science, with successive and more sophisticated layers constructed upon one another, such as, for instance, is illustrated in the diagram below.

$$\begin{array}{c}\text{Set Theory}\\ \downarrow \\ \text{Binary Relations}\\ \downarrow \\ \text{Order}\\ \downarrow \\ \text{Algebra}\\ \downarrow \\ \text{Topology}\\ \downarrow \\ \text{Functional Analysis etc}\end{array}$$

Traditionally, see also Sections 1.1 and 1.2, the problem of solving linear and nonlinear PDEs, is formulated in the context of the most sophisticated two levels, namely, that of topology and functional analysis. As a consequence of the almost exclusive use of the highly specialized tools of functional analysis, the basic underlying concepts involved in solving PDEs are only perceived through some of their most complicated aspects.

Furthermore, the Order Completion Method, in contradistinction with the usual methods discussed in Section 1.2, applies to situations which are far more general than PDEs alone. Indeed, the method is based on general results on the construction of Dedekind order complete partially ordered sets, and the extension of suitable mappings between such ordered sets. This is exactly the reason for its type independent power.

Now, as mentioned, the Order Completion Method goes far beyond the usual methods of functional analysis when it comes to the existence of generalized solutions to both linear and nonlinear PDEs. What is more, this method also produces a blanket regularity result for the solutions constructed. Furthermore, in view of the intuitively clear nature of the concept of order, one also gains much insight into the mechanisms involved in the solution of PDEs, as well as the structure of such generalized solutions. In this regard, the deescalation from the level of topology and functional analysis to the level of order proves to be particularly useful and relevant. However, when it comes to further properties of the solutions, such as for instance regularity of the solutions, functional analysis, or for that matter any mathematics, may yet play an important, but secondary role.

As mentioned, the theory of Order Completion is based on rather basic constructions in partially ordered sets. In this regard, let $(X, \leq_X)$ and $(Y, \leq_Y)$ be partially ordered sets. A mapping

$$T : X \ni x \mapsto Tx \in Y \tag{1.94}$$

which is *injective*, but not necessarily *surjective*, is an *order isomorphic embedding* whenever

$$\begin{array}{ll} \forall & x_0, x_1 \in X : \\ & Tx_0 \leq_Y Tx_1 \Leftrightarrow x_0 \leq_X x_1 \end{array} \tag{1.95}$$

A partially ordered set is called Dedekind Order Complete [101] if every set $A \subset X$ which is bounded from above has a least upper bound, and every set which is bounded from below has largest lower bound. That is,

$$\begin{array}{ll} \forall & A \subseteq X : \\ & 1) \left(\begin{array}{ll} \exists & x_0 \in X : \\ & x \leq_X x_0,\, x \in A \end{array} \right) \Rightarrow \left(\begin{array}{ll} \exists & u_0 \in X : \\ & (x \leq_X x_0,\, x \in A) \Rightarrow u_0 \leq_X x_0 \end{array} \right) \\ & 2) \left(\begin{array}{ll} \exists & x_0 \in X : \\ & x_0 \leq_X x,\, x \in A \end{array} \right) \Rightarrow \left(\begin{array}{ll} \exists & u_0 \in X : \\ & (x_0 \leq_X x,\, x \in A) \Rightarrow x_0 \leq_X u_0 \end{array} \right) \end{array} \tag{1.96}$$

With every partially ordered set $(X, \leq_X)$ one may associate a Dedekind Order Complete set $\left(X^\sharp, \leq_{X^\sharp}\right)$ and an order isomorphic embedding $\iota_X : X \to X^\sharp$, see [101], [102] and [119, Appendix A], such that

$$\begin{array}{ll} \forall & x^\sharp \in X^\sharp : \\ & 1) \;\; L\left(x^\sharp\right) = \{x \in X \,:\, \iota_X x \leq_{X^\sharp} x^\sharp\} \neq \emptyset \\ & 2) \;\; U\left(x^\sharp\right) = \{x \in X \,:\, x^\sharp \leq_{X^\sharp} \iota_X x\} \neq \emptyset \\ & 3) \;\; \sup L\left(x^\sharp\right) = \inf U\left(x^\sharp\right) = x^\sharp \end{array} . \tag{1.97}$$

Furthermore, if $\left(X_0^\sharp, \leq_{X_0^\sharp}\right)$ is another partially ordered set, and $\iota_{X,0} : X \to X_0^\sharp$ an order isomorphic embedding that satisfies (1.97), then there is a bijective order isomorphic embedding

$$T : X^\sharp \to X_0^\sharp,$$

or shortly an order isomorphism, so that the diagram

$$\begin{tikzcd} X^\sharp \arrow[rr, "T"] & & X_0^\sharp \\ & X \arrow[ul, "\iota_X"] \arrow[ur, "\iota_{X,0}"'] & \end{tikzcd} \tag{1.98}$$

commutes. That is, the Dedekind Order Completion of a partially ordered set is unique up to order isomorphism.

Also, see [119, Appendix A], if $T : X \to Y$ is an order isomorphic embedding, then T extends uniquely to an order isomorphic embedding

$$T^\sharp : X^\sharp \to Y^\sharp$$

so that the diagram

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[d, "\iota_X"'] & Y \arrow[d, "\iota_Y"] \\
X^\sharp \arrow[r, "T^\sharp"] & Y^\sharp
\end{tikzcd}$$

commutes. Moreover, for any $y_0 \in Y$, we have

$$\left(\begin{array}{ll} \exists ! & x^\sharp \in X^\sharp : \\ & T^\sharp x^\sharp = y_0 \end{array} \right) \Leftrightarrow \left(\begin{array}{ll} \exists & A \subseteq X : \\ & y_0 = \sup\{Tx \, : \, x \in A\} \end{array} \right) \tag{1.99}$$

It is within this general context of partially ordered sets and order isomorphic embeddings that the Order Completion Method [119] for nonlinear PDEs is formulated.

These basic results on the completion of partially ordered sets and the extensions of order isomorphisms, may be applied to large classes of nonlinear PDEs in so far as the existence, uniqueness and basic regularity of generalized solutions are concerned. In this regard we consider a general, nonlinear PDE

$$T(x, D)\, u(x) = f(x), \; x \in \Omega \text{ nonempty and open} \tag{1.100}$$

of order $m \geq 1$ arbitrary but fixed. The right hand term $f : \Omega \to \mathbb{R}$ is assumed to be continuous, and the partial differential operator $T(x, D)$ is supposed to be defined through a jointly continuous mapping

$$F : \Omega \times \mathbb{R}^M \to \mathbb{R} \tag{1.101}$$

by the expression

$$T(x, D)\, u(x) = F\left(x, u(x), ..., D^p u(x), ..\right), \; x \in \Omega \tag{1.102}$$

where $p \in \mathbb{N}^n$ satisfies $|p| \leq m$. As is well known, see for instance Sections 1.1 and 1.2, a general nonlinear PDE of the form (1.100) through (1.102) will in general fail to have classical solutions on the whole domain of definition Ω. However, a necessary condition for the existence of a solution to (1.100) on a neighborhood of any $x_0 \in \Omega$ may be formulated in terms of the mapping (1.101) in a rather straight forward way.

In this regard, suppose that for some $x_0 \in \Omega$ there is a neighborhood V of x_0 and a function function $u \in \mathcal{C}^m(V)$ that satisfies (1.100). That is,

$$\begin{array}{ll} \exists & V \in \mathcal{V}_{x_0} \text{ nonempty, open :} \\ \exists & u \in \mathcal{C}^m(\Omega) \ : \\ & T(x,D)\,u(x) = f(x)\,,\ x \in V \end{array} , \tag{1.103}$$

where $\mathcal{V}_{x_0}$ denotes set of open neighborhoods of x_0. Then, in view of (1.103), it is clear that

$$\begin{array}{ll} \forall & x \in V : \\ \exists & (\xi_p(x))_{|p| \leq m} \in \mathbb{R}^M : \\ & F(x, ..., \xi_p(x), ...) = f(x) \end{array} . \tag{1.104}$$

Now, in view of (1.103) and (1.104) it is clear that the condition

$$\begin{array}{ll} \forall & x \in \Omega : \\ & f(x) \in \left\{ F(x,\xi) \middle| \ \xi = (\xi_p)_{|p| \leq m} \in \mathbb{R}^M \right\} \end{array} \tag{1.105}$$

is nothing but a necessary condition for the existence of a classical solution $u \in \mathcal{C}^m(\Omega)$ to (1.100). Many PDEs of applicative interest satisfies (1.105) trivially. Indeed, in this regard we may note that, since the mapping F is continuous, the set

$$\left\{ F(x,\xi) \middle| \ \xi = (\xi_p)_{|p| \leq m} \in \mathbb{R}^M \right\}$$

must be an interval in $\mathbb{R}$ which is either bounded, half bounded or equals all of $\mathbb{R}$. In the particular case of a linear PDE

$$T(x,D)\,u(x) = \sum_{|\alpha| \leq m} a_\alpha(x)\, D^\alpha u(x)\,,\ x \in \Omega$$

that satisfies

$$\begin{array}{ll} \forall & x \in \Omega : \\ \exists & \alpha \in \mathbb{N}^n,\ |\alpha| \leq m : \\ & a_\alpha(x) \neq 0 \end{array}$$

we have

$$\left\{ F(x,\xi) \middle| \ \xi = (\xi_p)_{|p| \leq m} \in \mathbb{R}^M \right\} = \mathbb{R}. \tag{1.106}$$

This is also the case for most nonlinear PDEs of applicative interest, including large classes of polynomial nonlinear PDEs. As a mere *technical convenience*, we shall assume the slightly stronger condition

$$\begin{array}{ll} \forall & x \in \Omega : \\ & f(x) \in \operatorname{int}\left\{ F(x,\xi) \,\middle|\, \xi = (\xi_p)_{|p| \leq m} \in \mathbb{R}^M \right\} \end{array} \tag{1.107}$$

In view of (1.106) it follows that (1.107) is also satisfied by the respective classes of linear and nonlinear PDEs.

Subject to the assumption (1.107) one obtains the local approximation result

$$\begin{array}{ll} \forall & x_0 \in \Omega : \\ \forall & \epsilon > 0 : \\ \exists & \delta > 0 : \\ \exists & u \in \mathcal{C}^\infty(\Omega) : \\ & \|x - x_0\| \leq \delta \Rightarrow f(x) - \epsilon \leq T(x,D)\,u(x) \leq f(x) \end{array} \tag{1.108}$$

Indeed, from (1.107) it follows that, for $\epsilon > 0$ small enough, there is some $\xi^\epsilon \in \mathbb{R}^M$ so that

$$F(x_0, \xi^\epsilon) = f(x_0) - \frac{\epsilon}{2}. \tag{1.109}$$

Choosing $u \in \mathcal{C}^\infty(\Omega)$ in such a way that

$$\begin{array}{ll} \forall & |p| \leq m : \\ & D^p u(x_0) = \xi_p^\epsilon \end{array},$$

the result follows.

At this junction we should note that, in contradistinction with rather difficult techniques of the usual functional analytic methods, the local approximation condition (1.108) follows by from basic properties of continuous functions on Euclidean space. Furthermore, a *global* version of (1.108) is obtained as a straightforward application of (1.108) and the existence of a suitable tiling of Ω by compact sets, see for instance [58]. In this regard, we have

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & \Gamma_\epsilon \subset \Omega \text{ closed nowhere dense} : \\ \exists & u_\epsilon \in \mathcal{C}^\infty(\Omega \setminus \Gamma_\epsilon) : \\ & x \in \Omega \setminus \Gamma_\epsilon \Rightarrow f(x) - \epsilon \leq T(x,D)\,u_\epsilon(x) \leq f(x) \end{array} \tag{1.110}$$

The singularity set Γ_ϵ in (1.110) is typically generated as the union of countably many hyperplanes. As such, each of the singularity sets Γ_ϵ has zero Lebesgue measure, that is, $\operatorname{mes}(\Gamma_\epsilon) = 0$.

From a topological point of view, the approximation result (1.110) is extraordinarily versatile. Indeed, since the singularity set Γ_ϵ, for $\epsilon > 0$, may be constructed so that $\text{mes}\,(\Gamma_\epsilon) = 0$, the sequence (u_n) of approximating functions, corresponding to the sequence $\left(\epsilon_n = \frac{1}{n}\right)$ of real numbers, satisfies

$$\begin{array}{ll} \exists & E \subseteq \Omega : \\ & 1) \quad \Omega \setminus E \text{ is of First Baire Category} \\ & 2) \quad \text{mes}\,(\Omega \setminus E) = 0 \\ & 3) \quad x \in E \Rightarrow T\,(x,D)\,u_n\,(x) \to f\,(x) \end{array} \tag{1.111}$$

Over and above the mere *pointwise convergence almost everywhere*, the sequence (u_n) also satisfies the stronger condition

$$\begin{array}{ll} \exists & E \subseteq \Omega : \\ & 1) \quad \Omega \setminus E \text{ is of First Baire Category} \\ & 2) \quad \text{mes}\,(\Omega \setminus E) = 0 \\ & 3) \quad (T\,(x,D)\,u_n) \text{ converges to } f \text{ uniformly on } E \end{array} \tag{1.112}$$

Furthermore, since the singularity set $\Gamma_{\frac{1}{n}}$ associated with each function u_n in the sequence (u_n) is of measure 0, each such function is measurable on Ω. Moreover, one may construct each of the functions u_n in such a way that

$$\begin{array}{ll} \forall & K \subset \Omega \text{ compact :} \\ & \int_K |u_n\,(x)\,|dx < \infty \end{array} \tag{1.113}$$

Then, in view of (1.112) and (1.113) it follows by rather elementary arguments in measure theory that

$$\begin{array}{ll} \forall & K \subset \Omega \text{ compact :} \\ & \int_K |T\,(x,D)\,u_n\,(x) - f\,(x)\,|dx \to 0 \end{array} \tag{1.114}$$

The Order Completion Method, as developed in [119] and presented here, operates on a far more basic level than the respective topological interpretations (1.111) to (1.114) of the Global Approximation Result (1.110). In this regard, and in view of the closed nowhere dense singularity sets Γ_ϵ associated with the approximations (1.110), the family of spaces

$$\mathcal{C}^m_{nd}\,(\Omega) = \left\{ u : \Omega \to \mathbb{R} \;\middle|\; \begin{array}{ll} \exists & \Gamma_u \subset \Omega \text{ closed nowhere dense :} \\ & u \in \mathcal{C}^m\,(\Omega \setminus \Gamma_u) \end{array} \right\} \tag{1.115}$$

prove to be particularly useful. Indeed, in view of the continuity of the mapping F in (1.102) one may associate with the nonlinear partial differential operator $T\,(x,D)$ a mapping

$$T : \mathcal{C}^m_{nd}\,(\Omega) \to \mathcal{C}^0_{nd}\,(\Omega) \tag{1.116}$$

with the property

$$\begin{array}{ll} \forall & u \in \mathcal{C}^m_{nd}\,(\Omega)\;: \\ \forall & \Gamma \subset \Omega \text{ closed nowhere dense :} \\ & u \in \mathcal{C}^m\,(\Omega \setminus \Gamma) \Rightarrow Tu \in \mathcal{C}^0\,(\Omega \setminus \Gamma) \end{array}$$

The space $\mathcal{C}^0_{nd}(\Omega)$ contains certain pathologies. In particular, consider the simple example when $\Omega = \mathbb{R}$. For different values of $\alpha \in \mathbb{R}$, the functions

$$u_\alpha : \mathbb{R} \ni x \to \begin{cases} 0 & \text{if} \quad x \neq \alpha \\ 1 & \text{if} \quad x = \alpha \end{cases}$$

each corresponds to a different elements in $\mathcal{C}^0_{nd}(\Omega)$, although any two such functions are the same on some open and dense set. In order to remedy this apparent flaw, one may introduce an equivalence relation on the space $\mathcal{C}^0_{nd}(\Omega)$ through

$$u \sim v \Leftrightarrow \left(\begin{array}{ll} \exists & \Gamma \subset \Omega \text{ closed nowhere dense :} \\ & 1) \quad u, v \in \mathcal{C}^0(\Omega \setminus \Gamma) \\ & 2) \quad x \in \Gamma \Rightarrow u(x) = v(x) \end{array} \right) \tag{1.117}$$

The quotient space $\mathcal{C}^0_{nd}(\Omega) / \sim$ is denoted $\mathcal{M}^0(\Omega)$, and a partial order is defined on it through

$$U \leq V \Leftrightarrow \left(\begin{array}{ll} \exists & u \in U,\, v \in V : \\ \exists & \Gamma \subset \Omega \text{ closed nowhere dense :} \\ & 1) \quad u, v \in \mathcal{C}^0(\Omega \setminus \Gamma) \\ & 2) \quad x \in \Omega \setminus \Gamma \Rightarrow u(x) \leq v(x) \end{array} \right) \tag{1.118}$$

It should be noted that with the order (1.118) the space $\mathcal{M}^0(\Omega)$ is a lattice. Since

$$\begin{array}{ll} \forall & u, v \in \mathcal{C}^m_{nd}(\Omega) \subset \mathcal{C}^0_{nd}(\Omega) \ : \\ & u \sim v \Rightarrow Tu \sim Tv \end{array}$$

one may associate with the mapping (1.4) in a canonical way a mapping

$$\overline{T} : \mathcal{C}^m_{nd}(\Omega) \to \mathcal{M}^0(\Omega) \tag{1.119}$$

so that the diagram

$$\begin{tikzcd} \mathcal{C}^m_{nd}(\Omega) \arrow[rr, "T"] \arrow[dr, "\overline{T}"'] & & \mathcal{C}^0_{nd}(\Omega) \arrow[dl, "q_\sim"] \\ & \mathcal{M}^0(\Omega) & \end{tikzcd} \tag{1.120}$$

commutes, with $q_\sim$ the quotient map associated with the equivalence relation (1.117).

The equation

$$\overline{T}u = f, \tag{1.121}$$

which is a generalization of the PDE (1.100), does not fit in the framework (1.100) through (1.94). In particular, the mapping $\overline{T}$ is, except in extremely particular cases, neither injective, nor is it monotone. One may overcome these difficulties by associating an equivalence relation with the mapping $\overline{T}$ through

$$\begin{array}{ll} \forall & u, v \in \mathcal{C}^m_{nd}(\Omega) \ : \\ & u \sim_{\overline{T}} v \Leftrightarrow \overline{T}u = \overline{T}v \end{array} \tag{1.122}$$

We denote the quotient space $\mathcal{C}^m_{nd}(\Omega) / \sim_{\overline{T}}$ by $\mathcal{M}^m_{\overline{T}}(\Omega)$, and order it through

$$\begin{array}{ll} \forall & U, V \in \mathcal{M}^m_{\overline{T}}(\Omega) \ : \\ & U \leq_{\overline{T}} V \Leftrightarrow \left(\begin{array}{ll} \forall & u \in U,\ v \in V \ : \\ & \overline{T}u \leq \overline{T}v \end{array} \right) \end{array} \tag{1.123}$$

Clearly, one may now obtain, in a canonical way, an *injective* mapping

$$\widehat{T} : \mathcal{M}^m_{\overline{T}}(\Omega) \rightarrow \mathcal{M}^0(\Omega) \tag{1.124}$$

so that the diagram

$$\begin{tikzcd} \mathcal{C}^m_{nd}(\Omega) \arrow[r, "T"] \arrow[d, "q_{\overline{T}}"'] & \mathcal{C}^0_{nd}(\Omega) \arrow[d, "q_{\sim}"] \\ \mathcal{ML}^m_{\overline{T}}(\Omega) \arrow[r, "\widehat{T}"] & \mathcal{M}^0(\Omega) \end{tikzcd} \tag{1.125}$$

commutes, with $q_{\overline{T}}$ the quotient mapping associated with the equivalence relation (1.122). In particular, the mapping $\widehat{T}$ may be defined as

$$\begin{array}{ll} \forall & U \in \mathcal{M}^m_{\overline{T}}(\Omega) \ : \\ & \widehat{T} : U \mapsto \overline{T}u,\ u \in U \end{array} \tag{1.126}$$

The mapping $\widehat{T}$ clearly also satisfies (1.95) so that it is an order isomorphic embedding, and the setup (1.95) through (1.99) applies to the *generalized equation*

$$\widehat{T}U = f,\ U \in \mathcal{M}^m_{\overline{T}}(\Omega) \tag{1.127}$$

which, in view of the commutative diagrams (1.120) and (1.125), is equivalent to (1.121). In this regard, applying (1.110) and (1.99) leads to the fundamental *existence* and *uniqueness* result

$$\begin{array}{ll} \forall & f \in \mathcal{C}^0(\Omega) \text{ that satisfies (1.107) :} \\ \exists ! & U^\sharp \in \mathcal{M}_T^m(\Omega)^\sharp \; : \\ & \widehat{T}^\sharp U^\sharp = f \end{array} \tag{1.128}$$

where $\mathcal{M}_T^m(\Omega)^\sharp$ and $\mathcal{M}^0(\Omega)^\sharp$ are the Dedekind completions of $\mathcal{M}_T^m(\Omega)$ and $\mathcal{M}^0(\Omega)$, respectively, and

$$\widehat{T}^\sharp : \mathcal{M}_T^m(\Omega)^\sharp \to \mathcal{M}^0(\Omega)^\sharp$$

is the unique extension of $\widehat{T}$.

It should be noted that, in contradistinction with the usual functional analytic and topological methods, the arguments leading to the above existence and uniqueness result are particularly clear and transparent. Moreover, these arguments remain valid in a far more general setting. Furthermore, and as we have mentioned, in view of the *transparent* and *intuitively clear* way in which the Dedekind order completion of a partially ordered set is constructed, the structure of the generalized solution $U^\sharp \in \mathcal{M}_T^m(\Omega)^\sharp$ to (1.100) is particularly clear.

In this regard, we recall the abstract construction of the completion of a partially ordered set [101], [102], [119]. Consider a nonvoid partially ordered set $(X, \leq_X)$ which, for convenience, we assume is without top or bottom. Two dual operations on the powerset $\mathcal{P}(X)$ of X may be defined through

$$\mathcal{P}(X) \ni A \mapsto A^u = \bigcap_{a \in A} [a\rangle \in \mathcal{P}(X)$$

and

$$\mathcal{P}(X) \ni A \mapsto A^l = \bigcap_{a \in A} \langle a] \in \mathcal{P}(X)$$

where $[a\rangle$ and $\langle a]$ are the half bounded intervals

$$[a\rangle = \{x \in X \, : \, a \leq_X x\},$$

$$\langle a] = \{x \in X \, : \, x \leq_X a\}.$$

That is, A^u is the set of upper bounds of A, and A^l is the set of lower bounds of A. It is obvious that

$$A^u = X \Leftrightarrow A^l = X \Leftrightarrow A = \emptyset \tag{1.129}$$

while

$$A^u = \emptyset \Leftrightarrow A \text{ unbounded from above}$$

$$A^l = \emptyset \Leftrightarrow A \text{ unbounded from below.} \tag{1.130}$$

A set $A \in \mathcal{P}(X)$ is called a *cut* of the poset X if

$$A = A^{ul} = (A^u)^l.$$

The set of all cuts of X is denoted $\tilde{X}$, that is,

$$\tilde{X} = \{A \in \mathcal{P}(X) \,:\, A^{ul} = A\}.$$

In view of (1.129) and (1.130) we have

$$\emptyset, X \in \tilde{X}$$

so that $\tilde{X} \neq \emptyset$. Furthermore, for each $A \in \mathcal{P}(X)$ we have

$$A \subset A^{ul} \in \tilde{X} \tag{1.131}$$

and

$$\left(A^{ul}\right)^{ul} = A^{ul} \tag{1.132}$$

so that, in view of (1.131) and (1.132) we have

$$\tilde{X} = \{A^{ul} \,:\, A \in \mathcal{P}(X)\}.$$

In particular, for each $x \in X$ the set $\langle x]$ belongs to $\tilde{X}$. Furthermore, it is clear that

$$\begin{array}{ll} \forall & x_0, x_1 \in X : \\ & \langle x_0] \subseteq \langle x_1] \Leftrightarrow x_0 \leq x_1 \end{array}$$

so that the mapping

$$\iota_X : X \ni x \mapsto \langle x] \in \tilde{X} \tag{1.133}$$

is an order isomorphic embedding when $\tilde{X}$ is ordered through inclusion

$$A \leq B \Leftrightarrow A \subseteq B.$$

The main theorem in this regard, due to MacNeille, is the following.

Theorem 3 **[102] Let X be a partially ordered set without top or bottom, and $\tilde{X}$ the set of cuts in X, ordered through inclusion. Then the following statements are true:*

1. *The poset* $\left(\tilde{X}, \leq\right)$ *is order complete.*

2. *The order isomorphic embedding* ι_X *in (1.133) preserves infima and suprema, that is,*

$$\begin{array}{l} \forall \quad A \subseteq X : \\ \quad 1) \quad x_0 = \sup A \Rightarrow \iota_X x_0 = \sup \iota_X (A) \\ \quad 2) \quad x_0 = \inf A \Rightarrow \iota_X x_0 = \inf \iota_X (A) \end{array}$$

3. *For* $A \in \tilde{X}$*, we have the order density property* X *in* $\tilde{X}$*, namely*

$$\begin{aligned} A &= \sup_{\tilde{X}} \{\iota_X x : x \in X , \langle x] \subseteq A\} \\ &= \inf_{\tilde{X}} \{\iota_X x : x \in X , \langle x] \supseteq A\} \end{aligned}$$

An easy corollary to MacNeille's Theorem 3 is the following Dedekind Completion Theorem.

Corollary 4 *Suppose that the partially ordered set* X *is a lattice. Then the partially ordered set*

$$X^\sharp = \tilde{X} \setminus \{X, \emptyset\}$$

ordered through inclusion is a Dedekind complete lattice. Furthermore, the mapping

$$\iota_X : X \ni x \mapsto \langle x] \in X^\sharp$$

is an order isomorphic embedding which preserves infima and suprema. Furthermore, the order density property

$$\begin{aligned} A &= \sup_{\tilde{X}} \{\iota_X x : x \in X , \langle x] \subseteq A\} \\ &= \inf_{\tilde{X}} \{\iota_X x : x \in X , \langle x] \supseteq A\} \end{aligned}$$

also holds.

Now, in view of the above general construction, we may interpret the existence and uniqueness result (1.128) for the solutions to continuous nonlinear PDEs of the form (1.100) through (1.102), subject to the assumption (1.107) as follows. From the approximation result (1.110) and the definition (1.126) it follows that

$$f = \sup\{\widehat{T}U : U \in \mathcal{M}_T^m(\Omega), \widehat{T}U \leq f\}$$

Then, in view of the definition (1.123) of the partial order on $\mathcal{M}_T^m(\Omega)$, as well as Corollary 4, the generalized solution $U^\sharp$ to (1.100) may be expressed as

$$U^\sharp = \{U \in \mathcal{M}_T^m(\Omega) : \widehat{T}U \leq f\}$$

Finally, recalling the structure of the quotient space, and in particular the equivalence relation (1.122) one has

$$U^{\sharp} = \{u \in \mathcal{C}^m_{nd}(\Omega) \ : \ \overline{T}u \leq f\}$$

The unique generalized solution to (1.100) may therefore be interpreted as the totality of all subsolutions $u \in \mathcal{C}^m_{nd}(\Omega)$ to (1.100) which includes also all *exact classical* solutions, whenever such solutions exist, and all generalized solutions in $\mathcal{C}^m_{nd}(\Omega)$. In this regard, we may notice that the notion of generalized solution through Order Completion is *consistent* with the usual classical solutions as well as with generalized solutions in $\mathcal{C}^m_{nd}(\Omega)$.

The Order Completion Method also provides a *blanket regularity*, see [8] and [9], for the solutions to nonlinear PDEs, in the following sense, which is a consequence of the fact that the Dedekind completion of the space $\mathcal{M}^0(\Omega)$ may be represented as the set $\mathbb{H}_{nf}(\Omega)$ of all nearly finite Hausdorff continuous interval valued functions. In this regard, there is then an order isomorphism

$$F_0 : \mathcal{M}^0(\Omega)^{\sharp} \rightarrow \mathbb{H}_{nf}(\Omega) \tag{1.134}$$

Then, since the mapping

$$\widehat{T}^{\sharp} : \mathcal{M}^m_T(\Omega) \rightarrow \mathcal{M}^0(\Omega)^{\sharp}$$

is an order isomorphic embedding, it follows that

$$\widehat{T}^{\sharp} \circ F_0 : \mathcal{M}^m_T(\Omega) \rightarrow \mathbb{H}_{nf}(\Omega)$$

is an order isomorphic embedding. In this way, the generalized solution to (1.100) may be seen as being *assimilated* with usual Hausdorff continuous functions.

At this point we have only considered existence and uniqueness of generalized solutions to free problems. That is, we have only solved the equation (1.100) without imposing any addition boundary and / or initial conditions. It is well know that, in the traditional functional analytic approaches to PDEs, in particular those that involve weak solutions and distributions, the further problem of satisfying such additional conditions presents difficulties that most often require entirely new and rather difficult techniques. This is particularly true when distributions, their restrictions to lower dimensional manifolds and the associated trace operators are involved. In contrast to the well known difficulties caused by the presence of boundary and / or initial conditions in the customary methods for PDEs, the Order Completion Method incorporates such conditions in a rather straight forward and easy way, as demonstrated by several examples, see for instance [119, Part II]. This is achieved by first obtaining an appropriate version of the global approximation result (1.110) that incorporates the respective boundary and / or initial value problem. The key is what amounts to a separation of the problem of satisfying the PDE and that of satisfying the additional condition. In this way, boundary and / or initial value problems are solved, essentially, by the same techniques that apply to the free problem.

As a final remark concerning the theory for the existence and regularity of the solutions to arbitrary continuous nonlinear PDEs, as we have sketched it in this section, we mention certain possibilities for further enrichment of the basic theory. In particular, the following may serve as guidelines for such an enrichment.

(A) The space of generalized solutions to (1.100) may depend on the PDE operator $T(x, D)$

(B) There is no differential structure on the space of generalized solutions

In order to accommodate (A), one may do away with the equivalence relation (1.122) on $\mathcal{C}^m_{nd}(\Omega)$, and instead consider

$$u \sim v \Leftrightarrow \left(\begin{array}{ll} \exists & \Gamma \subset \Omega \text{ closed nowhere dense :} \\ & 1) \quad u, v \in \mathcal{C}^m(\Omega \setminus \Gamma) \\ & 2) \quad x \in \Omega \setminus \Gamma \Rightarrow u(x) = v(x) \end{array} \right)$$

to obtain the quotient space $\mathcal{M}^m(\Omega)$. Furthermore, one may consider a partial order other than (1.123), which does not depend on the partial differential operator $T(x, D)$. Indeed, in the original spirit of Sobolev spaces, one may consider the partial order

$$\begin{array}{ll} \forall & U, V \in \mathcal{M}^m(\Omega) \ : \\ & U \leq_D V \Leftrightarrow \left(\begin{array}{ll} \forall & |\alpha| \leq m : \\ & D^\alpha U \leq D^\alpha V \end{array} \right) \end{array}$$

which could also solve (B). However, such an approach presents several difficulties. In particular, the existence of generalized solutions in the Dedekind completion of the partially ordered set $(\mathcal{M}^m(\Omega), \leq_D)$ is not clear. In fact, the possibly nonlinear mapping T associated with the PDE (1.100) cannot be extended to the Dedekind completion in a unique and meaningful way, unless T satisfies some additional and rather restrictive conditions. We mention that the use of partial orders other than (1.123) was investigated in [119, Section 13], but the partial orders that are considered are still in some relation to the PDE operator $T(x, D)$. Regarding (B), we may recall that there is in general no connection between the usual order on $\mathcal{M}^m(\Omega) \subset \mathcal{M}^0(\Omega)$ and the derivatives of the functions that are its elements.

One possible way of going beyond the basic theory of Order Completion is motivated by the fact that the process of taking the supremum of a subset A of a partially ordered set X is essentially a process of approximation. Indeed,

$$x_0 = \sup A \tag{1.135}$$

means that the set A approximates x_0 arbitrarily close from below. Approximation, however, is essentially a topological process. In this regard, the various topological interpretations (1.111) through (1.114) of the global approximation result (1.110) present a myriad of new opportunities. Therefore, and in connection with (1.135), perhaps the most basic approach, and the one nearest the basic Theory of Order Completion would comprise a topological type model for the process of Dedekind completion of $\mathcal{M}^0(\Omega)$.

1.5 Beyond distributions

Ever since Schwartz [145] proved the so called Schwartz Impossibility Result, and Lewy [97] gave an example of a linear, variable coefficient PDE with no distributional solution, and in particular over the past four decades, there has been an increasing awareness that the usual methods of linear functional analysis, which are quite effective in the case of linear PDEs, in particular those with constant coefficients, cannot be reproduced in any general and consistent way when dealing with *nonlinear* PDEs.

As such, a number of other methods, mostly based on linear functional analysis, were introduced. Many of these theories, it should be mentioned, depend far less on the sophisticated tools of functional analysis than is the case, for instance, with the $\mathcal{D}$' distributions. In this regard, we mention here the Theory of Monotone Operators [31], and the theory of Viscosity Solutions, see for instance [42]. These methods, however, were developed with particular types of equations in mind, and their powers are therefore limited to those particular types of equations for which they were designed. It should be noted that, in those cases when these methods do apply, they have proven to be effective beyond the earlier functional analytic methods.

Recently, a general and type independent theory for the existence and regularity of generalized solutions of systems of nonlinear PDEs, based on techniques from Hilbert space, was initiated by J Neuberger, see [115] through [118]. This theory is based on a generalized method Steepest Descent in suitably constructed Hilbert spaces. Since this theory is not restricted to any particular class of nonlinear PDEs, that is, it is general and type independent, it bears comparison with methods developed in Part II of this work. As such, we include below a short account of the underlying ideas involved.

In this regard, let H be a suitable Hilbert space of generalized functions, for instance, H might be one of the Sobolev spaces $H^{2,m}(\Omega)$. For a given nonlinear PDE

$$T(x, D) u(x) = f(x), \tag{1.136}$$

the method is supposed to produce a generalized solution in H. In order to obtain such a generalized solution, a suitable real valued mapping

$$\phi_T : H \to \mathbb{R} \tag{1.137}$$

is associated with with the nonlinear partial differential operator $T(x, D)$ such that the *critical points* $u \in H$ of ϕ_T correspond to the solutions of (1.136) in H.

In this regard, recall that the derivative of a mapping

$$\phi : H \to \mathbb{R}$$

at $u \in H$ is a continuous linear mapping

$$\phi'_u : H \to \mathbb{R}$$

that satisfies

$$\lim_{v \to 0} \frac{|\phi(u+v) - \phi(u) - \phi'_u(v)|}{\|v\|} = 0.$$

We may associated with the function ϕ a mapping

$$D\phi : H \ni u \mapsto \phi'_u \in H' \tag{1.138}$$

where H' is the dual of the Hilbert space H. The mapping ϕ is $\mathcal{C}^1$-smooth on H whenever the mapping (1.138) is continuous. For such a $\mathcal{C}^1$-smooth mapping ϕ, the *gradient* of ϕ is the mapping

$$\nabla\phi : H \to H$$

such that

$$\forall \quad \begin{array}{l} u, v \in H : \\ \phi'_u(v) = \langle v, \nabla\phi(u)\rangle_H \end{array} . \tag{1.139}$$

The gradient mapping $\nabla\phi$ exists since $D\phi$ is continuous. A critical point of ϕ is any $u \in H$ such that $(\nabla\phi)(u) = 0$.

Neuberger's method involves techniques to show the existence of critical points to mappings (1.137) associated with a nonlinear PDE, as well as effective numerical computation of such critical points. This involves, inter alia, the adaptation of the gradient mapping (1.139), as well as modifications of Newton's Method of Steepest Descent to the particular problem at hand.

It should be noted that the underlying ideas upon which these methods are based do not depend on any particular form of the nonlinear partial differential operator $T(x, D)$. As such, the theory is, to a great extent, general and type independent. However, the relevant techniques involve several highly technical aspects, which have, as of yet, not been resolved for a class of equations comparable to that to which the Order Completion Method applies. On the other hand, the numerical computation of solutions, based on this theory, has advances beyond the scope of analytic techniques. In this regard, remarkable results have been obtained, see for instance [118].

Chapter 2

Topological Structures in Analysis

2.1 Point-Set Topology: From Hausdorff to Bourbaki

Topology, generally speaking, may be described as that part of mathematics that deals with shape and nearness without explicit reference to magnitudes. The first results of a topological nature date back to Euler, who solved the now well known 'Bridges of Köningstad' problem. Cantor, however, gave the first description of the topology on $\mathbb{R}$ in the modern spirit of the subject. Namely, Cantor introduced the concept of an open set in $\mathbb{R}$. It was only in 1906 when a general framework was introduced in which to describe such concepts as distance, nearness, neighborhood and convergence in an abstract setting.

In this regard Fréchet [59] introduced the concept of a metric space, which generalizes the Euclidean spaces

$$\mathbb{R}^n = \left\{ x = (x_i)_{i \leq n} \middle| \begin{array}{l} \forall \quad i \in \mathbb{N},\, i \leq n : \\ x_i \in \mathbb{R}^n \end{array} \right\},\ n \geq 1 \tag{2.1}$$

with the usual Euclidean metric

$$d_2 : \mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto \left(\sum_{i \leq n} (x_i - y_i)^2 \right)^{\frac{1}{2}} \in \mathbb{R}^+ \cup \{0\} \tag{2.2}$$

which, in the case $n = 3$, coincides with our everyday experience of the distance between two points in space. The concept of a metric space is a generalization of (2.1) to (2.2) in two different, yet equally important ways.

In the first place, the set $\mathbb{R}^n$ of n-tuples of real numbers is replaced by an arbitrary, nonempty set X. Furthermore, the mapping $d_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^+ \cup \{0\}$ is replaced by a suitable real valued mapping

$$d_X : X \times X \to \mathbb{R}^+ \cup \{0\} \tag{2.3}$$

where for each $x, y \in X$, the real number $d_X(x, y)$ is interpreted as the distance between x and y. The properties that the metric d_X in (2.3) is supposed to satisfy are suggested by our everyday experience with the distance between two points in three dimensional space, which may be expressed in mathematical form through (2.2). In particular, any two points in space are at a fixed, nonnegative distance from each other, which is expressed as the relation

$$\begin{array}{ll} \forall & x, y \in X : \\ & d_X(x, y) \geq 0 \end{array} . \tag{2.4}$$

Furthermore, the distance between any two distinct points is positive, that is,

$$\begin{array}{ll} \forall & x, y \in X : \\ & d(x, y) = 0 \Leftrightarrow x = y \end{array} . \tag{2.5}$$

Moreover, distance, as we commonly experience it, is a symmetric relation. That is, moving from point A to point B in three dimensional space along a straight line, one traverses the same distance as if we were moving from B to A. In general, this may be expressed as

$$\begin{array}{ll} \forall & x, y \in X : \\ & d_X(x, y) = d_X(y, x) \end{array} . \tag{2.6}$$

The fourth and final condition has rather strong geometric antecedents, as well as consequences. In this regard, consider two distinct points A and B in three dimensional space. The shortest path that can be traced out by a particle moving from the point A to the point B is the straight line connecting A and B. This may be generalized by the condition

$$\begin{array}{ll} \forall & x, y, z \in X : \\ & d_X(x, y) \leq d_X(x, z) + d_X(z, y) \end{array} . \tag{2.7}$$

Within such a general framework it is possible to describe a variety of concepts that are of importance in analysis. In this regard, one may formulate the notion of convergence of a sequence without reference to the particular properties of the elements of the underlying set X. Furthermore, Cantor's concept of an open set in $\mathbb{R}$ has a natural and straight forward generalization to metric spaces. Moreover, within the setting of metric spaces, one may define the fundamental concept of continuity of a function

$$u : X \to Y$$

in far more general cases than had previously been considered.

However, the way in which metric spaces are defined in (2.3) to (2.7) places rather serious limitations on the possible structures that may be conceived of in this way. A most simple example will serve to illustrate the limitations of the concept of a metric space.

Example 5 *Consider the set $\mathbb{R}^{\mathbb{R}}$ of all functions $u : \mathbb{R} \to \mathbb{R}$. A sequence (u_n) in $\mathbb{R}^{\mathbb{R}}$ converges pointwise to $u \in \mathbb{R}^{\mathbb{R}}$ whenever*

$$\begin{array}{ll} \forall & x \in \mathbb{R} : \\ \forall & \epsilon > 0 : \\ \exists & N(x,\epsilon) \in \mathbb{N} : \\ & n \geq N(x,\epsilon) \Rightarrow |u(x) - u_n(x)| < \epsilon \end{array} . \tag{2.8}$$

There is no metric d on $\mathbb{R}^{\mathbb{R}}$ for which (2.8) is equivalent to

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & N(\epsilon) \in \mathbb{N} : \\ & n \geq N(\epsilon) \Rightarrow d(u, u_n) < \epsilon \end{array} . \tag{2.9}$$

In this regard consider the function

$$u : \mathbb{R} \ni x \mapsto \begin{cases} 0 & if \quad x \notin \mathbb{Q} \\ 1 & if \quad x \in \mathbb{Q} \end{cases}$$

Write the set $\mathbb{Q}$ in the form of a sequence $\mathbb{Q} = \{q_1, q_2, ...\}$. For each $n \in \mathbb{N}$ let $\mathbb{Q}_n = \{q_1, ..., q_n\}$. For each $n \in \mathbb{N}$, define the function u_n as

$$u_n : \mathbb{R} \ni x \mapsto \begin{cases} 0 & if \quad x \notin \mathbb{Q}_n \\ 1 & if \quad x \in \mathbb{Q}_n \end{cases}$$

Clearly the sequence (u_n) converges pointwise to u.

Let $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ denote the subspace of $\mathbb{R}^{\mathbb{R}}$ consisting of all continuous functions $u : \mathbb{R} \to \mathbb{R}$. Suppose that there exists a metric d on $\mathbb{R}^{\mathbb{R}}$ so that (2.9) is equivalent to (2.8). Let $\mathcal{C}^0(\mathbb{R})^c$ denote the closure of $\mathcal{C}^0(\mathbb{R})$ in $\mathbb{R}^{\mathbb{R}}$ with respect to the metric d, that is,

$$\mathcal{C}^0(\mathbb{R})^c = \left\{ u \in \mathbb{R}^{\mathbb{R}} \,\middle|\, \begin{array}{ll} \exists & (u_n) \subset \mathcal{C}^0(\mathbb{R}) : \\ & (u_n) \text{ converges to } u \text{ w.r.t. } d \end{array} \right\}$$

Each function u_n is the pointwise limit of a sequence of continuous functions. As such, $u_n \in \mathcal{C}^0(\mathbb{R})^c$ for each $n \in \mathbb{N}$, and consequently $u \in \mathcal{C}^0(\mathbb{R})^c$. Therefore there exists a sequence (v_n) of continuous functions that converges pointwise to u. Then u is continuous everywhere except on a set of First Baire Category [121], which is clearly not the case.

In view of Example 5 it is clear that, even though Fréchet's theory of metric spaces provides a useful framework in which to formulate some problems in analysis, there are certain rather basic and important situations that cannot be described in terms of metric spaces. Furthermore, the axioms (2.4) through (2.7) of a metric are all based on our geometric intuition, an intuition which is most typically based solely on our experience of the usual three dimensional space we believe ourselves to reside in. As such, it turns out to fail in capturing other, more general notions of space.

In order to obtain more general structures than metric spaces, one may replace the conditions (2.3) to (2.7) on the mapping $d : X \times X \to \mathbb{R}$ that defines the metric space structure with less stringent ones to obtain, for instance, a pseudo-metric space. This, however, would miss the essence of the matter.

F Hausdorff was the first to free topology from our geometric intuition. In particular, the fundamental concept in Hausdorff's view of topology is not that of 'distance', but rather a generalization of Cantor's definition of open set. Moreover, this is done in purely set theoretic terms, and, in general, is does not involve any notion of magnitude.

In this regard, we recall what it means for two metrics d_0 and d_1 on a set X to be equivalent. Namely, d_0 and d_1 are considered equivalent whenever

$$\begin{array}{ll} \forall & x_0 \in X : \\ \exists & \alpha_{x_0,0}, \alpha_{x_0,1} > 0 : \\ & \alpha_{x_0,0} d_0 (x_0, x) \leq d_1 (x_0, x) \leq \alpha_{x_0,1} d_0 (x_0, x),\ x \in X \end{array}. \tag{2.10}$$

That is, every open ball with respect to d_0 contains an open ball with respect to d_1 and, conversely, every open ball with respect to d_1 contains an open ball with respect to d_0. Then it is clear that, for any two metrics d_0 and d_1 on X we have

$$\left(\begin{array}{ll} \forall & U \subseteq X : \\ & U \text{ open w.r.t. } d_0 \Leftrightarrow U \text{ open w.r.t. } d_1 \end{array} \right) \Leftrightarrow d_0 \text{ equivalent to } d_1 \tag{2.11}$$

In view of the equivalence of (2.10) through (2.11), we may describe a metric space (X, d) uniquely by either specifying the metric explicitly, or by specifying the collection τ of open sets. Therefore, in order to generalize the concept of metric space, one may, as mentioned, relax some of the conditions (2.4) to (2.7), or consider a suitable family of mappings. On the other hand, one may generalize the concept of 'open set'. Not only is the structure of the metric space uniquely and equivalently specified in terms of either one of these two concepts, but, and equally importantly, the continuous functions

$$u : X \to Y,$$

where Y is another metric space, are also uniquely and equivalently specified.

As mentioned, Hausdorff's concept of topology is based on a generalization of the concept of an open set. In this regard, one may deduce, from the definition of an open set in a metric space, and the axioms of a metric (2.3), certain properties of open sets which may be stated *without reference to the metric*. First of all, for some metric space (X, d_X), we note that, both the entire space X and the empty set $\emptyset$ are open. That is,

$$X, \emptyset \in \tau_{X_d}, \tag{2.12}$$

where τ_{X_d} denotes the collection of open subsets of X, with respect to the metric d_X. Indeed,

$$\begin{array}{ll} \forall & x_0 \in X : \\ \forall & \delta > 0 : \\ & B_\delta(x_0) \subseteq X \end{array}$$

and

$$\left\{ x_0 \in \emptyset \;\middle|\; \begin{array}{ll} \forall & \delta > 0 : \\ \exists & x \notin \emptyset : \\ & d_X(x_0, x) < 0 \end{array} \right\} = \emptyset$$

Furthermore, the collection of open sets are closed under the formation of finite intersections. That is,

$$\begin{array}{ll} \forall & U_1, ..., U_n \in \tau_{d_X} : \\ & U = \bigcap_{i=1}^n U_i \in \tau_{d_X} \end{array} \tag{2.13}$$

Indeed, is U either empty, or nonempty. In case U is nonempty, we have

$$\begin{array}{ll} \forall & x_0 \in U : \\ \forall & i = 1, ..., n : \\ \exists & \delta_{x_0,i} > 0 : \\ & B_{\delta_{x_0,i}}(x_0) \subseteq U_n \end{array}$$

Hence, upon setting $\delta_{x_0} = \inf\{\delta_{x_0,i} : i = 1, ..., n\}$, it is clear that $B_\delta(x_0) \subseteq U$, so that $U \in \tau_{d_X}$. Moreover, the union of any collection of open sets is open, so that

$$\begin{array}{ll} \forall & \{U_i : i \in I\} \subseteq \tau_{d_X} : \\ & U = \bigcup_{i \in I} U_i \in \tau_{d_X} \end{array} \tag{2.14}$$

It is these three properties (2.12), (2.13) and (2.14), together with the so called Hausdorff property

$$\begin{array}{ll} \forall & x_0, x_1 \in X : \\ \exists & U_0, U_1 \text{ open sets} : \\ & 1) \quad U_0 \cap U_1 = \emptyset \\ & 2) \quad x_0 \in U_0,\ x_1 \in U_1 \end{array} \tag{2.15}$$

that Hausdorff [70] set down as the axioms of his topology. In particular, Hausdorff defined a topology on a set X to be any collection τ of subsets of X that satisfies

(2.12), (2.13), (2.14) and (2.15), and termed the pair (X, τ) a topological space. With the minor revision of omitting the Hausdorff Axiom (2.15), this definition has remained unchanged for nearly a century. Hausdorff's theory was subsequently developed by several authors, chief among these being Kuratowski [93], [94], and the Bourbaki group [30].

In particular, the Bourbaki group, and most notably A Weil [160], introduced the highly important concept of a uniform space as a generalization of that of a metric space, within the context of topological spaces. The concept of uniform space allows for the definition of Cauchy sequences, or more generally Cauchy filters, as well as the associated concepts of completeness and completion.

Hausdorff's concept of topology, although rather abstract, proved extraordinarily useful, in particular in analysis. By the middle of the 20th century, mathematicians realized that Hausdorff's topology could provide the framework within which Banach's powerful results on normed vector spaces [17] could be generalized. With the subsequent development of the theory of locally convex spaces, the much sought generalizations were, to a limited extent, fulfilled. However, even at this early stage of development of topology, and its applications to mathematics in general, in particular analysis, certain deficiencies in general topology became apparent.

2.2 The Deficiencies of General Topology

As mentioned, Hausdorff's concept of topology proved to be particularly useful in generalizing classical result in analysis, for instance the powerful tools of linear functional analysis developed by Banach [17] within the setting of metric linear spaces. In spite of the great utility of these techniques, several serious deficiencies of General Topology had emerged by the middle of the twentieth century. In particular, the most important failure of the category of topological spaces is that it is not Cartesian closed, which is as much as to say that there is no natural topological structure for function spaces.

In this regard, recall that if X, Y and Z are sets, then one has the relation

$$Z^{X\times Y} \simeq \left(Z^X\right)^Y. \tag{2.16}$$

That is, there is a canonical one-to-one correspondence between functions

$$f : X \times Y \to Z \tag{2.17}$$

and functions

$$g : Y \to Z^X = \{h : X \to Z\}. \tag{2.18}$$

Indeed, with any function (2.17) we may associate the function

$$\tilde{f} : Y \ni y \to f(\cdot, y) \in Z^X. \tag{2.19}$$

That is,

$$\tilde{f}(y) : X \ni x \mapsto f(x,y) \in Z \tag{2.20}$$

Conversely, with a function (2.18) we can associate the mapping

$$\underline{g} : X \times Y \ni (x,y) \mapsto g(y)(x) \in Z \tag{2.21}$$

Within the context of topological spaces (2.16) is naturally formulated in terms of continuous functions. In this case, the exponential law (2.16) may be expressed as

$$\mathcal{C}(X \times Y, Z) \simeq \mathcal{C}(Y, \mathcal{C}(X,Z)), \tag{2.22}$$

In general, (2.22) is not satisfied. That is, there are plenty of topological spaces X, Y and Z, that are of significant interest, such that there is no topology τ on the spaces of continuous functions in (2.22) for which (2.22) holds.

Indeed, let

$$f : X \times Y \to Z. \tag{2.23}$$

be a continuous map. With the mapping (2.23) we may associate a mapping

$$F_f : Y \ni y \mapsto F_f(y) \in \mathcal{C}(X,Z) \tag{2.24}$$

defined as

$$F_f : X \ni x \mapsto f(x,y) \in Z$$

Conversely, with a mapping

$$F : Y \to \mathcal{C}(X,Z) \tag{2.25}$$

we may associate a mapping

$$f_F : X \times Y \to Z \tag{2.26}$$

defined as

$$f_F : X \times Y \ni (x,y) \mapsto (F(y))(x) \in Z$$

Suppose that $\mathcal{C}(X,Y)$ is equipped with the compact-open topology, which is specified by the subbasis

$$\left\{ S(K,U) \middle| \begin{array}{ll} 1) & K \subseteq X \text{ compact} \\ 2) & U \subseteq Y \text{ open} \end{array} \right\}$$

where

$$S(K,U) = \{f \in \mathcal{C}(X,Y) \ : \ f(K) \subseteq U\}.$$

In view of the continuity of the mapping (2.23), it follows that the associated mapping (2.24) must also be continuous. Conversely, if X is *locally compact* and *Hausdorff*, then the mapping (2.26) associated with the mapping (2.25) is continuous whenever the mapping (2.25) is continuous. Therefore, in case X is locally compact and Hausdorff, (2.23) through (2.26) specifies a bijective mapping

$$\chi : \mathcal{C}(X \times Z, Y) \ni f \mapsto F_f \in \mathcal{C}(Z, \mathcal{C}(X,Y)) \tag{2.27}$$

Moreover, if Y and Z are also locally compact, the mapping (2.27) is a homeomorphism, which is as much as to say that the exponential law (2.22) holds for locally compact spaces X, Y and Z, and the compact open topology on the relevant spaces of continuous functions.

This is known as the universal property of the compact open topology, within the class of locally compact spaces. However, when the assumption of local compactness on any of the spaces X, Y or Z is relaxed, then, in general, either the mapping (2.26) associated with (2.25) fails to be continuous, or the mapping (2.27) is no longer a homeomorphism, see for instance [110]. In particular, unless all the spaces X, Y and Z are locally compact, there is no topology on $\mathcal{C}(X,Y)$ so that the above construction holds.

Other rather unsatisfactory consequence of the mentioned categorical failure of topology, namely, that the category of topological spaces is not Cartesian closed, appears in connection with quotient mappings. Recall that a topological quotient map is a surjective map

$$q : X \to Y$$

between topological space X and Y that satisfies

$$\begin{array}{ll} \forall & U \subseteq Y : \\ & U \text{ open in } Y \ \Leftrightarrow q^{-1}(U) \text{ open in } X \end{array} \tag{2.28}$$

Such maps appear frequently in topology, as well as in its applications to analysis. However, as mentioned, several irregularities appear in connection with quotient mappings [11], of which we mention only the following.

Quotient maps are not hereditary. That is, if $q : X \to Y$ is a quotient mapping, and A a subset of Y, then the surjective, continuous mapping

$$q_A : q^{-1}(A) \to A$$

obtained by restriction q to $q^{-1}(A)$ in X, is, in general, not a quotient map with respect to the subspace topologies on A and $q^{-1}(A)$. To see that this is so, consider the following example [126].

Example 6 *Consider the sets* $X = \{0, 1, 2\}$ *and* $Y = \{0, 1, 2, 3\}$*. On* X *consider the topology* $\tau_X = \{\{0, 2\}, \{1, 3\}, X, \emptyset\}$*, and equip* Y *with the topology* $\tau_Y = \{Y, \emptyset\}$*. Then the mapping*

$$q : X \ni x \mapsto \begin{cases} 0 & if \quad x = 0 \\ 1 & if \quad x = 1 \\ 2 & if \quad x \in \{2, 3\} \end{cases}$$

is continuous and surjective. In particular, q *is a quotient map. Now consider the subset* $A = \{0, 1\}$ *of* Y*, so that* $B = q^{-1}(A) = \{0, 1\}$*. The subspace topology on* A *is* $\tau_A = \{A, \emptyset\}$*, while the subspace topology on* B *is* $\tau_B = \{\{0\}, \{1\}, B, \emptyset\}$*. Clearly the mapping* q_A*, which is simply the mapping* q *restricted to* B*, is not a quotient map.*

Furthermore, quotient maps are not productive. That is, if $\{X_i : i \in I\}$ and $\{Y_i : i \in I\}$ are families of topological spaces, and for each $i \in I$, the mapping

$$q_i : X_i \to Y_i$$

is a quotient map, then the product of the family of mappings $\{q_i : i \in I\}$, which is defined as

$$q : \prod_{i \in I} X_i \ni \mathbf{x} = (x_i)_{i \in I} \mapsto q(\mathbf{x}) = (q_i(x_i))_{i \in I} \in \prod_{i \in I} Y_i, \tag{2.29}$$

is not a quotient map with respect to the product topologies on $\prod X_i$ and $\prod Y_i$. Indeed, consider the following example [126].

Example 7 *On* $\mathbb{R}$ *consider the equivalence relation* $\sim$ *defined through*

$$\begin{array}{ll} \forall & x_0, x_1 \in \mathbb{R} : \\ & x_0 \sim x_1 \Leftrightarrow x_0 = x_1 \ or \ \{x_0, x_1\} \subset \mathbb{Z} \end{array} \tag{2.30}$$

Let $q_\sim : \mathbb{R} \to \mathbb{R}/\sim$ *denote the canonical map associated with the equivalence relation (2.30), and equip* $\mathbb{R}/\sim$ *with the quotient topology*

$$\begin{array}{ll} \forall & U \subseteq \mathbb{R}/\sim_\pi \ : \\ & U \in \tau_\omega \Leftrightarrow q_\sim^-(U) \ \ open \ in \ \mathbb{R} \end{array}$$

so that $q_\sim$ *is a quotient map. Let* $id_\mathbb{Q}$ *denote the identity mapping on the rational numbers* $\mathbb{Q}$*. Then the mapping*

$$q_\sim \times id_\mathbb{Q} : \mathbb{R} \times \mathbb{Q} \ni (x, r) \mapsto (q_\sim x, r) \in \mathbb{R} \times \mathbb{Q}$$

is continuous and surjective. If $q_\sim \times id_\mathbb{Q}$ *were a quotient map, then it would map saturated closed sets onto closed sets. Recall that a closed set is saturated if it is the inverse image of a subset of* $\mathbb{R}/\sim \times \mathbb{Q}$ *under* $q_\sim \times id_\mathbb{Q}$*. Let* (a_n) *be a sequence* *of irrational numbers converging to* 0. *For each* $n \in \mathbb{N}$, *let* $(r_{n,m})$ *be a sequence of rational numbers converging to* a_n. *Set*

$$A = \left\{ \left(n + \frac{1}{m}, r_{n,m} \right) \ : n, m \in \mathbb{N} \text{ and } m > 1 \right\}.$$

A *is closed and saturated in* $\mathbb{R} \times \mathbb{Q}$, *but* $(q_\sim \times id_\mathbb{Q})(A)$ *is not closed in* $\mathbb{R}/\sim \times \mathbb{Q}$.

As we have mentioned, the lack of a 'universal topological structure' for function spaces, as well as the well known difficulties that appear in connection with quotient mappings, is in fact only a concrete manifestation of the fundamental categorical flaw in Hausdorff's topology. Namely, that the category **TOP** of all topological spaces with continuous mappings is not Cartesian closed. This flaw of **TOP** manifests itself even in the relatively simple setting of locally convex linear topological spaces. In particular, if E is a locally convex space with topological dual E^*, then, unless E is a normable space, there is no locally convex topology on the dual E^* so that the simple evaluation mapping

$$ev : E \times E^* \ni (x, x^*) \mapsto x^*(x) \in \mathbb{K} \tag{2.31}$$

is continuous, with $\mathbb{K}$ the scalar field $\mathbb{R}$ or $\mathbb{C}$. To see that this is so, suppose that E is a locally convex space so that the evaluation mapping (2.31) is continuous with respect to some vector space topology $\tau_{\mathcal{L}E}$ on $\mathcal{L}E$. Since (2.31) is continuous at $(0,0)$, there is a zero neighborhood W in $\mathcal{L}E$, and a convex zero neighborhood U in E such that $W(U)$ is contained in the unit disc in $\mathbb{K}$. Since W is absorbing, U is bounded in the weak topology and therefore bounded in E. Since E contains a bounded zero neighborhood, it is normed.

In view of the remarks above, it is clear that Hausdorff's concept of 'topological space' is not a satisfactory one. In particular, due to the categorial failures of the category **TOP**, there is no natural structure on function spaces. One solution to this problem is provided by the theory of convergence spaces [26].

2.3 Convergence Spaces

As was observed in Section 2.2, the Hausdorff-Kuratowski-Bourbaki concept of topology suffers from serious deficiencies, which manifest themselves even in the relatively simple setting of locally convex topological vector spaces. Over and above these basic flaws, this notion of topology is rather restrictive, which may be seen from the fact that several useful modes of convergence cannot be adequately described in terms of the usual topology. In this regard, one may recall the following examples, see [120], [126] and [154].

Example 8 *Consider on the real line* $\mathbb{R}$ *the usual Lebesgue measure mes, and denote by* $M(\mathbb{R})$ *the space of all almost everywhere (a.e.) finite, measurable functions on*

$\mathbb{R}$, with the conventional identification of functions a.e. equal. A natural notion of convergence on $M(\mathbb{R})$ is that of convergence a.e. That is,

$$\begin{array}{l} \forall \quad (u_n) \subset M(\mathbb{R}) \ : \\ \forall \quad u \in M(\mathbb{R}) \ : \\ \quad (u_n) \ \text{converges a.e. to } u \Leftrightarrow \left(\begin{array}{ll} \exists & E \subset \mathbb{R},\ mes(E) = 0 : \\ & x \in \mathbb{R} \setminus E \Rightarrow u_n(x) \to u(x) \end{array} \right) \end{array} \tag{2.32}$$

There is no topology τ on $M(\mathbb{R})$ so that a sequence converges with respect to τ if and only if it converges a.e. to the same function. To see this, suppose that such a topology, say τ_{ae}, exists, and let (u_n) be a sequence which converges in measure to 0, but fails to converge a.e.. Then there is a τ_{ae} neighborhood V of the constant zero function, and a subsequence (u_{n_m}) of (u_n) so that

$$\begin{array}{ll} \forall & n_m \in \mathbb{N} : \\ & u_{n_m} \notin V \end{array} \tag{2.33}$$

Since (u_n) converges to 0 in measure, so does the subsequence (u_{n_m}). A well know theorem, see for instance [85], states that the subsequence (u_{n_m}) contains a further subsequence $\left(u_{n_{m_k}}\right)$ that converges a.e. to 0. Therefore the sequence $\left(u_{n_{m_k}}\right)$ is eventually in V, which contradicts (2.33).

Example 9 *Consider on the space $\mathcal{C}^0(\mathbb{R})$ of all continuous, real valued functions on $\mathbb{R}$ the pointwise order*

$$u \leq v \Leftrightarrow \left(\begin{array}{ll} \forall & x \in \mathbb{R} : \\ & u(x) \leq v(x) \end{array} \right) \tag{2.34}$$

With respect to this order, and the usual vector space operations, the space $\mathcal{C}(\mathbb{R})$ is an Archimedean vector lattice [101]. A sequence (u_n) in $\mathcal{C}^0(\mathbb{R})$ order converges to $u \in \mathcal{C}^0(\mathbb{R})$ whenever

$$\begin{array}{ll} \exists & (\lambda_n),\ (\mu_n) \subset \mathcal{C}^0(\mathbb{R}) \ : \\ & 1) \quad \lambda_n \leq \lambda_{n+1} \leq u_n \leq \mu_{n+1} \leq \mu_n,\ n \in \mathbb{N} \\ & 2) \quad \sup\{\lambda_n : n \in \mathbb{N}\} = u = \inf\{\mu_n : n \in \mathbb{N}\} \end{array} \tag{2.35}$$

There is no topology τ_o on $\mathcal{C}^0(\mathbb{R})$ so that a sequence converges to $u \in \mathcal{C}^0(\mathbb{R})$ with respect to τ_o if and only if it order converges to u. To see this, consider the sequence (u_n) defined as

$$u_n(x) = \begin{cases} 1 - n|x - q_n| & if \quad |x - q_n| < \frac{1}{n} \\ \\ 0 & if \quad |x - q_n| \geq \frac{1}{n} \end{cases}$$

Here $\mathbb{Q} \cap [0,1] = \{q_n : n \in \mathbb{N}\}$ is the set of rational numbers in the interval $[0,1]$, ordered as usual, so that the complement of any finite subset of $\mathbb{Q} \cap [0,1]$ is dense in $[0,1]$. For any $N_0 \in \mathbb{N}$, we have

$$\sup\{u_n : n \geq N_0\}(x) = 1,\ x \in \mathbb{R}$$

and

$$\inf\{u_n : n \geq N_0\}(x) = 0,\ x \in \mathbb{R}$$

so that this sequence does not order converge to 0. *Suppose that there exists a topology* τ_o *on* $\mathcal{C}^0(\mathbb{R})$ *that induces order convergence. Then there is some* τ_o*-neighborhood* V *of* 0, *and a subsequence* (u_{n_m}) *of* (u_n) *which is always outside of* V. *Let* (q_{n_m}) *denote the sequence of rational numbers associated with the subsequence* (u_{n_m}). *Since this sequence is bounded, there is a subsequence* $\left(q_{n_{m_k}}\right)$ *of it, and some real number* $q \in [0,1]$ *so that* $\left(q_{n_{m_k}}\right)$ *converges to* q. *Let* $\left(u_{n_{m_k}}\right)$ *be the subsequence of* (u_{n_m}) *corresponding to the sequence of rational numbers* $\left(q_{n_{m_k}}\right)$. *Then it is clear that*

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & N_\epsilon \in \mathbb{N} : \\ & n_{m_k} > N_\epsilon \Rightarrow u_{n_{m_k}}(x) = 0,\ |x-q| > \epsilon \end{array}$$

Set $\epsilon_l = \frac{1}{l}$ *for each* $l \in \mathbb{N}$. *Define the sequence* $\left(\mu_{n_{m_k}}\right)$ *in* $\mathcal{C}^0(\mathbb{R})$ *as*

$$\mu_{n_{m_k}}(x) = \begin{cases} 0 & if \quad |x-q| \geq 2\epsilon_l \\ 1 & if \quad |x-q| \leq \epsilon_l \\ \frac{|q-x|}{\epsilon_l} + 2 & if \quad \epsilon_l < |x-q| < 2\epsilon_l \end{cases}$$

whenever $N_{\epsilon_l} < n_{m_k} < N_{\epsilon_{l+1}}$. *The sequence* $\left(\mu_{n_{m_k}}\right)$ *decreases to* 0, *and*

$$u_{n_{m_k}} \leq \mu_{n_{m_k}},\ n_{m_k} \in \mathbb{N}$$

so that $\left(u_{n_{m_k}}\right)$ *order converges to* 0. *Therefore, it must eventually be in* V, *a contradiction. Therefore the topology* τ_o *cannot exists.*

We have given two useful and well known examples of concepts of sequential convergence that cannot be described in terms of the usual Hausdorff-Kuratowski-Bourbaki formulation of topology. Note that the concept of order convergence as introduced in Example 9 may be formulated in terms of an arbitrary partially ordered set $(X, \leq)$. In particular, if X is the set $M(\mathbb{R})$ of usual measurable functions on $\mathbb{R}$, modulo almost everywhere equal functions, with the pointwise a.e. order, then (2.32) is identical with the order convergence. The order convergence is widely used, particularly in the theory of vector lattices, where, for instance, it appears in connection with σ-continuous operators, and in particular integral operators [163].

In view of Examples 8 and 9, as well as the lack of a natural topological structure for function spaces discussed in Section 2.2, a more general notion of topology may be introduced. In this regard, we recall that a given topological space (X, τ) may be completely described by specifying the convergence associated with the topology

τ. More precisely, for each $x \in X$, we may specify the set $\lambda_\tau(x)$ consisting of those *filters* on X which converge to x with respect to τ. That is,

$$\lambda_\tau(x) = \{\mathcal{F} \text{ a filter on } X : \mathcal{V}_\tau(x) \subseteq \mathcal{F}\} \tag{2.36}$$

where $\mathcal{V}_\tau(x)$ denotes the τ-neighborhood filter at $x \in X$. In particular, if (x_n) is a sequence in X, then we may associate with it its Frechét filter

$$\langle (x_n) \rangle = [\{\{x_n : n \geq k\} : k \in \mathbb{N}\}].$$

If (x_n) converges to $x \in X$ with respect to the topology τ on X, that is,

$$\begin{array}{ll} \forall & V \in \mathcal{V}_\tau(X) : \\ \exists & N_V \in \mathbb{N} : \\ & n \geq N_V \Rightarrow x_n \in V \end{array}, \tag{2.37}$$

then we must have

$$\mathcal{V}_\tau(x) \subseteq \langle (x_n) \rangle$$

Conversely, if $\mathcal{V}_\tau(x) \subseteq \langle (x_n) \rangle$, then we must have (2.37). As such, the definition (2.36) of filter convergence in a topological space is nothing but a straight forward generalization of the corresponding notion for sequences.[1]

Remark 10 *Recall that a filter $\mathcal{F}$ on X is a nonempty collection of nonempty subsets of X such that*

$$\begin{array}{ll} \forall & F \in \mathcal{F} : \\ \forall & G \subseteq X : \\ & F \subseteq G \Rightarrow G \in \mathcal{F} \end{array}$$

and

$$\begin{array}{ll} \forall & F, G \in \mathcal{F} : \\ & F \cap G \in \mathcal{F} \end{array}.$$

A filter base for a filter is any collection $\mathcal{B} \subseteq \mathcal{F}$ so that

$$[\mathcal{B}] = \left\{ F \subseteq X \;\middle|\; \begin{array}{ll} \exists & B \in \mathcal{B} : \\ & B \subseteq F \end{array} \right\} = \mathcal{F}.$$

An ultrafilter on X is a filter which is not properly contained in any other filter. In particular, for each $x \in X$, the filter

$$[x] = \{F \subseteq X : x \in F\}$$

is an ultrafilter on X. The intersection of two filters $\mathcal{F}$ and $\mathcal{G}$ on X is defined as

$$\mathcal{F} \cap \mathcal{G} = \left\{ H \subseteq X \;\middle|\; \begin{array}{ll} \exists & F \in \mathcal{F}, G \in \mathcal{G} : \\ & F \cup G \subseteq H \end{array} \right\}$$

and it is the largest filter, with respect to inclusion, contained in both $\mathcal{G}$ and $\mathcal{F}$. A filter $\mathcal{F}$ is finer than $\mathcal{G}$, or alternatively $\mathcal{G}$ is coarser than $\mathcal{F}$, whenever $\mathcal{G} \subseteq \mathcal{F}$.

[1] In the sequel, we will make no distinction between a sequence and its associated Frechét filter. As such, we will denote both entities by (x_n). The meaning will be clear from the context.

More generally, a convergence structure [26] on a set X is defined as follows.

Definition 11 *A convergence structure on a nonempty set X is a mapping λ from X into the powerset of the set of all filters on X that, for each $x \in X$, satisfies the following properties.*

(i) $[x] \in \lambda(x)$

(ii) *If $\mathcal{F}$, $\mathcal{G} \in \lambda(x)$ then $\mathcal{F} \cap \mathcal{G} \in \lambda(x)$*

(iii) *If $\mathcal{F} \in \lambda(x)$ and $\mathcal{F} \subseteq \mathcal{G}$ then $\mathcal{G} \in \lambda(x)$.*

The pair (X, λ) is called a convergence space. When $\mathcal{F} \in \lambda(x)$ we say that $\mathcal{F}$ converges to x and write "$\mathcal{F} \to x$".

Concepts of convergences and convergence spaces that are more general than topological spaces were introduced and developed by several authors, see for instance [33], [37], [38], [51], [57], [78], [79] and [89]. The above Definition 11 is widely used and has proved to be a rather convenient one.

It is clear that the mapping λ_τ associated with a topology τ on a set X through (2.36) is a convergence structure. However, the concept of a convergence structure is far more general than that of a topology. Indeed, convergence almost everywhere of measurable functions, and the order convergence of sequences in $\mathcal{C}(\mathbb{R})$ discussed in Examples 8 and 9, respectively, are induced by suitable convergence structures, but cannot be induced by a topology. The most striking generalization inherent in Definition 11 may be formulated as follows. For every $x \in X$, the set of filters $\lambda_\tau(x)$ that converge to $x \in X$ with respect to the topology τ on X has a least element with respect to inclusion. Namely, the neighborhood filter $\mathcal{V}_\tau(x)$ at $x \in X$. In particular, for each $x \in X$ we have

$$\mathcal{V}_\tau(x) = \bigcap_{\mathcal{F} \in \lambda_\tau(x)} \mathcal{F} = \left\{ V \subseteq X \;\middle|\; \begin{array}{ll} \forall & \mathcal{F} \in \lambda_\tau(x) \ : \\ & V \in \mathcal{F} \end{array} \right\}. \tag{2.38}$$

More generally, for every subset $\{\mathcal{F}_i \, : \, i \in I\}$ of $\lambda_\tau(x)$, the filter

$$\bigcap_{i \in I} \mathcal{F}_i = \left[\left\{ F \subseteq X \;\middle|\; \begin{array}{ll} \forall & i \in I : \\ \exists & F_i \in \mathcal{F}_i : \\ & \cup_{i \in I} F_i \subseteq F \end{array} \right\} \right] \tag{2.39}$$

converges to x with respect to the topology τ. Clearly this need not be the case for a convergence space in general, see for instance [26]. However, topological concepts such as open set, closure of a set and continuity generalize to the more general context of convergence spaces in a natural way.

In this regard, if X and Y are convergence spaces with convergence structures λ_X and λ_Y, respectively, then a mapping

$$f : X \to Y$$

is *continuous at* $x \in X$ whenever

$$\mathcal{F} \in \lambda_X(x) \Rightarrow f(\mathcal{F}) = [\{f(F) \,:\, F \in \mathcal{F}\}] \in \lambda_Y(f(x))$$

and f is continuous on X if it is continuous at every point x of X. Furthermore, such a continuous mapping f is an *embedding* if it is injective with a continuous inverse defined on its image, and it is an *isomorphism* if it is also surjective.

The open subsets of a convergence space X are defined through the concept of *neighborhood*. Note that in the topological case, it follows from (2.38) that for any $x \in X$, a set $V \subseteq X$ is a neighborhood of x if and only if

$$\mathcal{F} \in \lambda_\tau(x) \Rightarrow V \in \mathcal{F}. \tag{2.40}$$

The definition of a neighborhood of a point x in an arbitrary convergence space is a straightforward generalization of (2.40). Namely,

$$V \in \mathcal{V}_{\lambda_X}(x) \Leftrightarrow \left(\begin{array}{cc} \forall & \mathcal{F} \in \lambda_X(x) \,: \\ & V \in \mathcal{F} \end{array} \right)$$

where $\mathcal{V}_{\lambda_X}(x)$ denotes the neighborhood filter at $x \in X$ with respect to the convergence structure λ_X.[2] A set $V \subseteq X$ is *open* if and only if it is a neighborhood of each of its elements.

The generalization of the closure of a subset A of a topological space X within the context of convergence spaces is the adherence. In the case of a topological space X, the closure of a subset A of X consists of all cluster points of A, that is,

$$\mathrm{cl}_\tau(A) = \left\{ x \in X \,\middle|\, \begin{array}{cc} \forall & V \in \mathcal{V}_x \,: \\ & V \cap A \neq \emptyset \end{array} \right\}. \tag{2.41}$$

Therefore, for each $x \in \mathrm{cl}(X)$, the filter

$$\mathcal{F} = [\{A \cap V \,:\, V \in \mathcal{V}_x\}]$$

converges to x, and $A \in \mathcal{F}$. Conversely, if there is a filter $\mathcal{F} \in \lambda_\tau(x)$ so that $A \in \mathcal{F}$, then in view of (2.36) it follows that A meets every neighborhood of x, so that $x \in \mathrm{cl}(A)$. That is, the closure of A consists of all points $x \in X$ so that A belongs to some filter $\mathcal{F} \in \lambda_\tau(x)$. The generalization of (2.41) to convergence spaces gives rise to the concept of adherence. The *adherence* of a subset A of a convergence space X is the set

$$a_{\lambda_X}(A) = \left\{ x \in X \,\middle|\, \begin{array}{cc} \exists & \mathcal{F} \in \lambda_X(x) \,: \\ & A \in \mathcal{F} \end{array} \right\}. \tag{2.42}$$

The set A is called *closed* if $a_{\lambda_X}(A) = A$.[3]

---

[2] If the convergence structure or topology is clear from the context, we will use the simplified notation $\mathcal{V}_x$ for the neighborhood filter at $x \in X$.

[3] Whenever there is no confusion, the adherence of a set A will simply be denoted by $a(A)$

Here we should point out that, although the concepts of open set, adherence and closed set coincide with the usual topological notions whenever the convergence space is topological, there are in general some important differences [26]. In particular, the *neighborhood filter* $\mathcal{V}_x$ at $x \in X$ need not converge to x, while the adherence operator will typically fail to be idempotent, that is,

$$a(A) \neq a(a(A))$$

A convergence space X that satisfies

$$\begin{array}{ll} \forall & x \in X : \\ & \mathcal{V}_x \in \lambda_X(x) \end{array}$$

is called *pretopological*, and the convergence structure λ_X is called a *pretopology* [26].

The customary constructions for producing new topological spaces form given ones, namely, initial and final structures, are defined for convergence spaces in the obvious way. Given a set X and a family of convergence structures $(X_i, \lambda_{X_i})_{i\in I}$ together with mappings

$$f_i : X \to X_i,\ i \in I \tag{2.43}$$

the *initial convergence structure* λ_X on X with respect to the family of mappings (2.43) is the coarsest convergence structure on X making each of the mappings $f_i : X \to X_i$ continuous. That is, for any other convergence structure λ on X such that all the f_i are continuous, we have

$$\lambda(x) \subseteq \lambda_X(x),\ x \in X. \tag{2.44}$$

The initial convergence structure is defined as

$$\mathcal{F} \in \lambda_X(x) \Leftrightarrow \left(\begin{array}{ll} \forall & i \in I : \\ & f_i(\mathcal{F}) \in \lambda_{X_i}(x) \end{array} \right)$$

Typical examples of initial convergence structures include the subspace convergence structure and the product convergence structure.

Example 12 *Let X be a convergence space, and A a subset of X. The subspace convergence structure λ_A induced on A from X is the initial convergence structure with respect to the inclusion mapping*

$$i_A : A \ni x \mapsto x \in X.$$

That is,

$$\mathcal{F} \in \lambda_A(x) \Leftrightarrow \left[\left\{ G \subseteq X \;\middle|\; \begin{array}{ll} \exists & F \in \mathcal{F} : \\ & F \subseteq G \end{array} \right\} \right] \in \lambda_X(x). \tag{2.45}$$

Example 13 *Consider a family $(X_i)_{i\in I}$ of convergence spaces, and let X be the Cartesian product of the family*

$$X=\prod_{i\in I}X_i.$$

The product convergence structure on X is the initial convergence structure with respect to the projection mappings

$$\pi_i : X\ni (x_i)_{i\in I}\mapsto x_i\in X_i,\ i\in I.$$

The convergent filters in the product convergence structure may be constructed as follows: A filter $\mathcal{F}$ on X converges to $x=(x_i)_{i\in I}\in X$ if and only if

$$\begin{array}{ll}\forall & i\in I :\\ \exists & \mathcal{F}_i\in\lambda_{X_i}(x_i)\ :\ .\\ & \prod_{i\in I}\mathcal{F}_i\subseteq\mathcal{F}\end{array}\tag{2.46}$$

Here $\prod_{i\in I}\mathcal{F}_i$ denotes the Tychonoff product of the filters $\mathcal{F}_i$, that is, $\prod_{i\in I}\mathcal{F}_i$ is the filter generated by

$$\left\{\prod_{i\in I}F_i\left|\begin{array}{l}F_i\in\mathcal{F}_i,\ i\in I\\ F_i\neq X_i\textit{ for only finitely many } i\in I\end{array}\right.\right\}.\tag{2.47}$$

Final structures are constructed in a similar way, and include quotient convergence structures and convergence inductive limits as particular cases. In this regard, given a set X, a family of convergence spaces $(X_i)_{i\in I}$, and mappings

$$f_i : X_i\rightarrow X,\ i\in I\tag{2.48}$$

the *final convergence structure* λ_X on X is the finest convergence structure making all the mappings (2.48) continuous. That is, for every convergence structure λ on X for which each mapping f_i is continuous, one has

$$\lambda_X(x)\subseteq\lambda(x),\ x\in X.\tag{2.49}$$

In particular, the final convergence structure on X is defined through

$$\lambda_X(x)=\{[x]\}\bigcup\left\{\mathcal{F}\left|\begin{array}{ll}\exists & i_1,...,i_k\in I :\\ \exists & x_n\in X_{i_n},\ n=1,...,k :\\ \exists & \mathcal{F}_n\in\lambda_{X_{i_n}}(x_n),\ n=1,...,k :\\ & \text{1)}\quad f_i(x_n)=x,\ i=1,...,k\\ & \text{2)}\quad f_{i_1}(\mathcal{F}_1)\cap...\cap f_{i_k}(\mathcal{F}_k)\subseteq\mathcal{F}\end{array}\right.\right\}.\tag{2.50}$$

Example 14 *Let X be a convergence space, Y a set and $q : X\rightarrow Y$ a surjective mapping. The quotient convergence structure on Y is the final convergence structure* *with respect to the mapping q. In particular, a filter $\mathcal{F}$ on Y converges to $y \in Y$ with respect to the quotient convergence structure λ_q on Y if and only if*

$$\begin{array}{ll} \exists & x_1, ..., x_k \in q^{-1}(y) \subseteq X : \\ \exists & \mathcal{F}_1, ..., \mathcal{F}_k \text{ filters on } X : \\ & \quad 1) \quad \mathcal{F}_i \in \lambda_X(x_i), \; i = 1, ..., k \\ & \quad 2) \quad q(\mathcal{F}_1) \cap ... \cap q(\mathcal{F}_k) \subseteq \mathcal{F} \end{array}. \tag{2.51}$$

If X and Y are convergence spaces, and $q : X \to Y$ a surjection so that Y carries the quotient convergence structure with respect to q, then q is called a convergence quotient mapping.

Remark 15 *In general, it is not true that a topological quotient mapping is a convergence quotient mapping. Indeed, if X and Y are topological spaces, and*

$$q : X \to Y$$

a continuous mapping, then q is a convergence quotient mapping if and only q is almost open [84], that is,

$$\begin{array}{ll} \forall & y \in Y : \\ \exists & x \in q^{-1}(y) : \\ \exists & \mathcal{B}_x \text{ a basis of open sets at } x : \\ & B \in \mathcal{B}_x \Rightarrow q(B) \text{ open in } Y \end{array}.$$

Within the category **CONV** of convergence spaces, the most striking deficiencies of topological spaces are resolved. In particular, in contradistinction with **TOP**, the category **CONV** is cartesian closed. As such, within this larger category, there is a natural *convergence structure* for function spaces, namely, the continuous convergence structure [26], [28].

If X and Y are convergence spaces, then the *continuous convergence structure* λ_c on the set $\mathcal{C}(X,Y)$ of continuous mappings from X into Y is the coarsest convergence structure making the evaluation mapping

$$\omega_{X,Y} : \mathcal{C}(X,Y) \times X \ni (f,x) \mapsto f(x) \in Y$$

continuous. That is, for each $f \in \mathcal{C}(X,Y)$ and every filter $\mathcal{H}$ on $\mathcal{C}(X,Y)$ we have

$$\mathcal{H} \in \lambda_c(f) \Leftrightarrow \left(\begin{array}{ll} \forall & x \in X : \\ \forall & \mathcal{F} \in \lambda_X(x) : \\ & \omega_{X,Y}(\mathcal{H} \times \mathcal{F}) \in \lambda_Y(f(x)) \end{array} \right). \tag{2.52}$$

For convergence spaces X, Y and Z, the mapping [4]

$$P : \mathcal{C}_c(X \times Y, Z) \to \mathcal{C}_c(X, \mathcal{C}_c(Y,Z))$$

[4]It is customary in the literature to denote by $\mathcal{C}_c(X,Y)$ the set of continuous functions from X into Y equipped with the continuous convergence structure.

which is defined as

$$P(f)(x) : Y \ni y \mapsto f(x, y) \in Z$$

is a homeomorphism, which shows that the category **CONV** is indeed cartesian closed.

Other difficulties encountered when working exclusively with topological spaces, such as for instance some of those mentioned in connection with quotient mappings, may also be resolved by considering the more general setting of convergence spaces. In this regard, we mention only the following.

Example 16 *In the category of convergence spaces quotient mappings are hereditary. Indeed, let X and Y be convergence spaces, and*

$$q : X \to Y$$

a surjective mapping so that Y carries the quotient convergence structure with respect to q. Consider any subspace A of Y, and the surjective mapping

$$q_A : q^{-1}(A) \ni x \mapsto q(x) \in A. \tag{2.53}$$

Clearly the subspace convergence structure on A is coarser than the quotient convergence structure induced by the mapping (2.53). Let the filter $\mathcal{F}$ on A converge to $y \in A$ with respect to the subspace convergence structure. That is,

$$\begin{array}{ll} \exists & \mathcal{F}_1, ..., \mathcal{F}_k \text{ filters on } X : \\ \exists & x_1, ..., x_k \in q^{-1}(y) : \\ & q(\mathcal{F}_1) \cap ... \cap q(\mathcal{F}_k) \subseteq \mathcal{F}_Y \end{array}$$

where $\mathcal{F}_Y$ denotes the filter generated by $\mathcal{F}$ in Y. Equivalently, we may say

$$\begin{array}{ll} \forall & F_1 \in \mathcal{F}_1, ..., F_k \in \mathcal{F}_k : \\ \exists & F \in \mathcal{F} : \\ & q(F_1) \cup ... \cup q(F_1) \supseteq F \end{array}. \tag{2.54}$$

We may assume that each filter $\mathcal{F}_i$ has a trace on $q^{-1}(A)$. That is,

$$\begin{array}{ll} \forall & F_i \in \mathcal{F}_i : \\ & F_i \cap q^{-1}(A) \neq \emptyset \end{array}.$$

As such, for each $i = 1, ..., k$ the filter $\mathcal{F}_{i|A}$ generated in $q^{-1}(A)$ by the family

$$\{F_i \cap q^{-1}(A) \ : \ F_i \in \mathcal{F}_i\}$$

converges to x_i in $q^{-1}(A)$. From (2.54) and the fact that $q^{-1}(A)$ is saturated with respect to q, it follows that

$$q_A(\mathcal{F}_{1|A}) \cap ... \cap q_A(\mathcal{F}_{i|A}) \subseteq \mathcal{F}$$

so that $\mathcal{F}$ converges to y with respect to the quotient convergence structure on A.

Example 17 *In the category of convergence spaces, quotient mappings are productive. That is, if* $(X_i)_{i\in I}$ *and* $(Y_i)_{i\in I}$ *are families of convergence spaces, and for each* $i \in I$ *the mapping*

$$q_i : X_i \to Y_i$$

is a convergence quotient mapping, then the surjective mapping

$$q : X = \prod_{i\in I} X_i \ni (x_i)_{i\in I} \mapsto (q_i x_i)_{i\in I} \in Y = \prod_{i\in I} Y_i$$

is a convergence quotient mapping. In this regard, and in view of (2.46) and (2.51) a filter $\mathcal{F}$ *on* Y *converges to* $y = (y_i)_{i\in I} \in Y$ *if and only if*

$$\begin{array}{ll} \forall & i \in I : \\ \exists & x_{i,1}, ..., x_{i,k_i} \in q_i^{-1}(y_i) : \\ \exists & \mathcal{F}_{i,1} \in \lambda_{X_i}(x_{i,1}), ..., \mathcal{F}_{i,k_i} \in \lambda_{X_i}(x_{i,k_i}) : \\ & \mathcal{F} \supseteq \prod_{i\in I} (q_i(\mathcal{F}_{i,1}) \cap ... \cap q_i(\mathcal{F}_{i,k_i})) \end{array} \tag{2.55}$$

where the product of filters in (2.55) is the Tychonoff product (2.47). An elementary, yet somewhat lengthy, computation shows that (2.55) coincides with the quotient convergence structure with respect to the mapping q*.*

The theory of convergence spaces has proven to be particularly powerful in so far as its applications to topology and analysis are concerned [26]. This effectiveness of convergence structures is due mainly to the fact that, as we have mentioned, the category of convergence spaces is cartesian closed, thus providing a suitable topological structure for function spaces. In this regard, we mention only the following.

The continuous convergence structure (2.52) yields a function space representation of a large class of topological and convergence spaces. In this regard, recall [26] that for a convergence space X, the mapping

$$i_X : X \to \mathcal{C}_c(\mathcal{C}_c(X)) \tag{2.56}$$

defined through

$$i_X(x) : \mathcal{C}_c(X) \ni f \mapsto f(x) \in \mathbb{R},$$

where $\mathcal{C}_c(X)$ is the set of real valued continuous functions on X equipped with the continuous convergence structure, is continuous. The convergence space X is called *c-embedded* whenever the mapping (2.56) is an embedding.

A characterization of c-embedded spaces may be obtained through the concepts of functionally regular and functionally Hausdorff convergence spaces. Recall [26] that a convergence space X is *functionally regular* if the initial topology τ on X with respect to $\mathcal{C}(X)$ is regular, and it is called *functionally Hausdorff* if τ is Hausdorff. The main result in this regard is the following.

Theorem 18 ** [26] A convergence space X is c-embedded if and only if X is functionally regular, functionally Hausdorff and Choquet [26]. That is, a filter $\mathcal{F}$ converges to $x \in X$ whenever every finer ultrafilter converges to x.*
In particular, $i_X(X) \subset \mathcal{C}_c(\mathcal{C}_c(X))$ is the set of all continuous algebra homomorphisms from $\mathcal{C}_c(X)$ into $\mathbb{R}$.

The class of c-embedded convergence spaces includes, amongst others, all Tychonoff spaces. However, not every c-embedded topological space is a Tychonoff space. Furthermore, products, subspaces and projective limits of c-embedded convergence spaces are again c-embedded.

Within the setting of functional analysis, in particular the theory of locally convex spaces, the theory of convergence spaces proves to be highly effective. In particular, the continuous convergence structure provides a natural structure for the topological dual of a locally convex space. In this regard, we may recall from Section 2.2 that for a locally convex space X, there is no locally convex topology on its dual $\mathcal{L}X$ so that the simple evaluation mapping

$$\omega_X : X \times \mathcal{L}X \ni (x, \varphi) \mapsto \varphi(x) \in \mathbb{K} \tag{2.57}$$

is continuous *unless X is normable.*

However, within the more general framework of convergence spaces, there is a natural dual structure available for locally convex spaces, namely, the continuous convergence structure. In this regard, we recall [26] that a convergence structure λ_X on a set X is a vector space convergence structure, and the pair (X, λ_X) a convergence space, if X is a vector space over some field of scalars $\mathbb{K}$, and the vector space operations

$$+ : X \times X \ni (x, y) \mapsto x + y \in X$$

and

$$\cdot : \mathbb{K} \times X \ni (\alpha, x) \mapsto \alpha x \in X$$

are continuous. For a convergence vector space X, we denote by $\mathcal{L}_c X$ the convergence vector space of all continuous linear functionals on X into $\mathbb{K}$ equipped with the continuous convergence structure. For any convergence vector space X, the space $\mathcal{L}_c X$ is a convergence vector space, and we call it the continuous dual of X. In this regard, the main result [26] is that the evaluation mapping

$$\omega_X : X \times \mathcal{L}_c X \to \mathbb{K}$$

is jointly continuous. Consequently, the natural mapping

$$i_X : X \to \mathcal{L}_c \mathcal{L}_c X \tag{2.58}$$

from X into its second dual, which is defined as

$$i_X(x) : \mathcal{L}_c X \ni \varphi \mapsto \varphi(x) \in \mathbb{K}, \tag{2.59}$$

is also continuous. In particular, in case X is a locally convex topological space, the mapping in (2.58) to (2.59) is actually an embedding. Furthermore, if X is complete, then this mapping is an isomorphism. Thus the continuous convergence structure provides a natural structure for the dual of a locally convex space.

Beyond the basic duality result for locally convex spaces, convergence vector spaces prove to be a far more natural setting for functional analysis, in comparison with locally convex spaces. This is so even if one's primary interest lies in the topological, locally convex case. In this regard, we may mention that the Pták's Closed Graph Theorem, a technical and notoriously difficult result in locally convex spaces, becomes a transparent and natural result when viewed in the setting of convergence vector spaces, see [21], [22], [25] and [26]. Furthermore, the scope of the Banach-Steinhauss Theorem is greatly expanded by formulating the problem in terms of convergence vector spaces [24]. Moreover, common and important objects such as the inductive limit of a family of locally convex spaces seem to be far removed from its component spaces, when viewed in the setting of locally convex spaces. Consequently, properties of the component spaces rarely translate to properties of the limit, while properties of the limit is not easily lifted to that of the component spaces, see for instance [27] for an indication of such difficulties. In contradistinction with the locally convex topological case, when such constructions are performed in the context of convergence vector spaces, there is a clear connection between components and limits.

Lastly, we mention that, owing to the remarkable categorical properties of convergence structures, the theory of convergence spaces has been applied with a good deal of success to difficult problems in point set-topology. In this regard, we mention the recent application to product theorems for topological spaces [51], [111], [112], [113].

2.4 Uniform Structures

We may recall from Section 2.1 that Hausdorff's topological spaces were introduced as a generalization of Frechét's metric spaces. Indeed, the concepts of open set, closed set, convergence of sequences, or more generally filters and nets, are extended in a straightforward way to this significantly more general class of spaces.

However, certain aspects of the structure of a metric space are not preserved in this generalization, namely, the uniform structure. In this regard, recall that for a metric space X with metric d_X, a sequence (x_n) on X is a *Cauchy sequence* if and only if

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & N_\epsilon \in \mathbb{N} : \\ & n, m \geq N_\epsilon \Rightarrow d_X(x_n, x_m) < \epsilon \end{array} . \tag{2.60}$$

Furthermore, if Y is a another metric space, then a $f : X \to Y$ is *uniformly continuous* whenever

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & \delta_\epsilon > 0 : \\ & d_X(x, y) < \delta_\epsilon \Rightarrow d_Y(f(x), f(y)) < \epsilon \end{array}$$

The space X is called *complete* if every Cauchy sequence in X converges to some $x \in X$. Moreover, with every metric space X one may associate a complete metric space $X^\sharp$, with metric $d_{X^\sharp}$, which is minimal in the following sense: There exists a uniformly continuous *embedding*

$$\iota_X : X \to X^\sharp$$

so that $\iota_X(X)$ is *dense* in $X^\sharp$. Furthermore, for any complete metric space Y, and any uniformly continuous mapping

$$f : X \to Y,$$

there exists a uniformly continuous mapping

$$f^\sharp : X^\sharp \to Y$$

so that the diagram

$$\begin{tikzcd} X \arrow[rr, "\iota_X"] \arrow[dr, "f"'] & & X^\sharp \arrow[dl, "f^\sharp"] \\ & Y & \end{tikzcd} \tag{2.61}$$

commutes. The above construction was given for the first time by Hausdorff [70]. The interest in such constructions may be seen from the fact that the set of real numbers may be constructed as the completion of a metric space, namely, the metric space of all rational numbers

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z} \right\}$$

with the usual metric

$$d(x, y) = |x - y|.$$

When one tries to extend the uniform concepts of Cauchy sequence, completeness and uniform continuity of functions to the more general setting of arbitrary topological spaces, there are several difficulties. These difficulties are all, to some extent, related to the following basic fact. Namely, that the uniform structure of a metric space, and in fact the metric itself, is defined through *binary relations*, whereas a topology is defined in a *unary* way. These underlying issues are identified within the most general context so far considered in [132], [133] and [134]. The structures introduced in [132] to [134], however, are far more general than the Hausdorff-Kuratowski-Bourbaki concept of topology, and contain it as a particular case.

As a further clarification of the issue raised in the previous paragraph, let us consider the more formal way to introduce the uniform structure of a metric space. In this regard, we may consider the family $\mathcal{U}_{X,d_X}$ of subsets of $X \times X$ specified through

$$U \in \mathcal{U}_{X,d_X} \Leftrightarrow \left(\begin{array}{ll} \exists & \epsilon > 0 : \\ & U_\epsilon \subseteq U \end{array} \right) \tag{2.62}$$

where, for $\epsilon > 0$, we define the set U_ϵ as

$$U_\epsilon = \{(x, y) \in X \times X \, : \, d_X(x, y) < \epsilon\}. \tag{2.63}$$

It is clear now that, for any sequence (x_n) on X, the sequence is a Cauchy sequence if and only if

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & N_\epsilon \in \mathbb{N} : \\ & n, m > N_\epsilon \Rightarrow \{(x_n, x_m) \, : \, n, m > N_\epsilon\} \subset U_\epsilon \end{array} . \tag{2.64}$$

Furthermore, for any other metric space Y, and any mapping $f : X \to Y$, the mapping f is uniformly continuous if and only if

$$\begin{array}{ll} \forall & U \in \mathcal{U}_{Y,d_Y} : \\ & (f^{-1} \times f^{-1})(U) = \{(x, y) \, : \, (f(x), f(y)) \in U\} \in \mathcal{U}_{X,d_X} \end{array} . \tag{2.65}$$

In view of (2.64) and (2.65) it is clear that, just as the open neighborhoods in X characterize the topology on X, the family of sets $\mathcal{U}_{X,d_X}$ *completely determines* the uniform structure of the metric space. In this regard, within the context of topological spaces, given a set X, the uniform structure shall consist of a family $\mathcal{U}_X$ of subsets of $X \times X$ that satisfy certain purely *set theoretic* properties, properties that are intended to capture suitable topological properties. The concept of a uniform space is then nothing but a distillation of these purely set-theoretic properties of $\mathcal{U}_X$. In particular, the Bourbaki group [30], and most notably A Weil [160], defined a *uniformity* on a set X as follows.

Definition 19 *A uniformity on a set X is a filter $\mathcal{U}$ on $X \times X$ that satisfies the following properties*

(i) *$\Delta \subseteq U$ for each $U \in \mathcal{U}$.*

(ii) *If $U \in \mathcal{U}$ then $U^{-1} \in \mathcal{U}$.*

(iii) *For each $U \in \mathcal{U}$ there is some $V \in \mathcal{U}$ so that $V \circ V \subseteq U$.*

Remark 20 *If U and V are subsets of the cartesian product $X \times X$ of X, then the inverse of U is defined as*

$$U^{-1} = \{(x,y) \in X \times X \,:\, (y,x) \in U\}$$

while the composition of U and V is specified through

$$U \circ V = \left\{ (x,y) \in X \times X \,\middle|\, \begin{array}{ll} \exists & z \in X : \\ & (x,z) \in V,\ (z,y) \in U \end{array} \right\},$$

Furthermore, for any set $A \subseteq X$, the set $U[A]$ is defined as

$$U[A] = \left\{ y \in X \,\middle|\, \begin{array}{ll} \exists & x \in A : \\ & (x,y) \in U \end{array} \right\}.$$

If A is the singleton $\{x\}$, then we simply write $U[x]$ for $U[A]$.

It is clear that the family (2.62) and (2.63) of subsets of the cartesian product $X \times X$ of a metric space X constitutes a uniformity. However, not every uniformity can be induced by a metric through (2.62) and (2.63). Conversely, every uniformity $\mathcal{U}_X$ on X induces a topology $\tau_{\mathcal{U}_X}$ on X. In particular, $\tau_{\mathcal{U}_X}$ may be defined as

$$W \in \tau_{\mathcal{U}_X} \Leftrightarrow \left(\begin{array}{ll} \forall & x \in W : \\ \exists & U \in \mathcal{U}_X : \\ & U[x] \subseteq W \end{array} \right). \tag{2.66}$$

Even though the topology (2.66) induced by a uniformity $\mathcal{U}_X$ on X can in general not be induced by a metric, the conditions (i) to (iii) in Definition 19 have strong metric antecedents [81]. Indeed, the first condition

$$\Delta \subseteq U,\, U \in \mathcal{U}_X$$

is derived from the property (2.5), while the second condition

$$U \in \mathcal{U}_X \Rightarrow U^{-1} \in \mathcal{U}_X$$

merely reflects the symmetry condition (2.6) of a metric. Lastly, the third condition

$$\begin{array}{ll} \forall & U \in \mathcal{U}_X : \\ \exists & V \in \mathcal{U}_X : \\ & V \circ V \subseteq U \end{array}$$

is an abstraction of the triangle inequality (2.7), and states, roughly speaking, that for ϵ-balls there are $\epsilon/2$-balls.

In view of (2.64) it is clear that a sequence (x_n) in a metric space X is a Cauchy sequence if and only if

$$\mathcal{U}_{X,d} \subseteq (x_n) \times (x_n),$$

where (x_n) denotes also the Fréchet filter associated with the sequence. As such, we may generalize the concept of a Cauchy sequence to any uniform space X as

$$(x_n) \text{ a Cauchy sequence } \Leftrightarrow \mathcal{U}_X \subseteq (x_n) \times (x_n).$$

More generally, given any filter $\mathcal{F}$ on X, we say that $\mathcal{F}$ is a *Cauchy filter* if and only if

$$\mathcal{U}_X \subseteq \mathcal{F} \times \mathcal{F}$$

Furthermore, a uniform space X is complete if and only if every Cauchy filter on X converges to some $x \in X$.

Other uniform concepts generalize in the expected way to uniform spaces. In this regard, we may recall [81] that a mapping

$$f : X \to Y,$$

with X and Y uniform spaces, is *uniformly continuous* if and only if

$$\forall \quad \begin{array}{l} U \in \mathcal{U}_Y : \\ (f^{-1} \times f^{-})(U) \in \mathcal{U}_X \end{array}.$$

Moreover, a uniformly continuous mapping is a *uniform embedding* whenever it is injective, and its inverse f^{-1} is uniformly continuous on the subspace $f(X)$ of Y. Furthermore, f is a uniform isomorphism is a uniform embedding which is also surjective.

The main result in connection with *completeness* of uniform spaces generalizes the corresponding result for metric spaces mentioned above, and is due to Weil [160]. However, it applies only to *Hausdorff* uniform spaces, that is, uniform spaces for which the induced topology (2.66) is Hausdorff. In this regard, for any Hausdorff uniform space X there is a complete, Hausdorff uniform space $X^\sharp$ and a uniform embedding

$$\iota_X : X \to X^\sharp$$

so that $\iota_X(X)$ is dense in $X^\sharp$. Furthermore, for any complete, Hausdorff uniform space Y, and any uniformly continuous mapping

$$f : X \to Y$$

there is a uniformly continuous mapping

$$f^\sharp : X^\sharp \to Y$$

so that the diagram

$$\begin{tikzcd} X \arrow[rr, "\iota_X"] \arrow[dr, "f"'] & & X^\sharp \arrow[dl, "f^\sharp"] \\ & Y & \end{tikzcd} \tag{2.67}$$

commutes.

It should be noted that not every topology τ_X on a set X can be induced by a uniformity through (2.66). Indeed, Weil [160] showed that for a given topology τ_X on X, there is a uniformity on X that induces τ_X through (2.66) if and only if τ_X is completely regular. As such, the class of uniform spaces is rather small in comparison with the class of all topological spaces. In this regard, several generalizations of a uniform space have been introduced in the literature, including that of a quasi-uniform space, which is related to the concept of nonsymmetric distance, see for instance [90].

Within the more general context of convergence spaces, a number of different concepts of 'uniform space' have been studied. The most successful of these are the so called uniform convergence spaces, and Cauchy spaces. The motivation for introducing such concepts within the setting of convergence spaces is twofold. First of all, it allows for the definition of uniform concepts, in particular that of completeness and completion, in this context, concepts which are fundamental in analysis. The second reason, and related to the first, is the mentioned relatively narrow applicability of uniform spaces within the context of the usual topology.

A uniform convergence space generalizes the concept of a uniform space in the following ways. Every uniformity on a set X gives rise to a uniform convergence structure. Furthermore, and as will be shown shortly, every uniform convergence structure induces a convergence structure. This induced convergence structure need not be, and in general is not, topological, and satisfies rather general separation properties. In particular, even in case the induced convergence structure is topological, it need not be completely regular. The definition of a uniform convergence space is now as follows [26].

Definition 21 *A uniform convergence structure on a set X is a family $\mathcal{J}_X$ on $X \times X$ that satisfies the following properties.*

(i) $[x] \times [x] \in \mathcal{J}_X$ *for every* $x \in X$[5].

(ii) *If* $\mathcal{U} \in \mathcal{J}_X$, *and* $\mathcal{U} \subseteq \mathcal{V}$, *then* $\mathcal{V} \in \mathcal{J}_X$.

(iii) *If* $\mathcal{U}, \mathcal{V} \in \mathcal{J}_X$, *then* $\mathcal{U} \cap \mathcal{V} \in \mathcal{J}_X$.

(iv) $\mathcal{U}^{-1} \in \mathcal{J}_X$ *whenever* $\mathcal{U} \in \mathcal{J}_X$.

(v) *If* $\mathcal{U}, \mathcal{V} \in \mathcal{J}_X$, *then* $\mathcal{U} \circ \mathcal{V} \in \mathcal{J}_X$ *whenever the composition exists.*

If $\mathcal{J}_X$ is a uniform convergence structure on X, then we refer to the pair $(X, \mathcal{J}_X)$ as a uniform convergence space.

Remark 22 *Let $\mathcal{U}$ and $\mathcal{V}$ filters on $X \times X$. The filter $\mathcal{U}^{-1}$ is defined as*

$$\mathcal{U}^{-1} = [\{U^{-1} : U \in \mathcal{U}\}].$$

If the filters $\mathcal{U}$ and $\mathcal{V}$ satisfies

$$\begin{array}{ll} \forall & U \in \mathcal{U} : \\ \forall & V \in \mathcal{V} : \\ & U \circ V \neq \emptyset \end{array},$$

then the filter $\mathcal{U} \circ \mathcal{V}$ exists, and is defined as

$$\mathcal{U} \circ \mathcal{V} = [\{U \circ V : U \in \mathcal{U},\, V \in \mathcal{V}\}].$$

As mentioned, every uniformity $\mathcal{J}_X$ on a set X induces a uniform convergence structure $\mathcal{J}_{\mathcal{U}_X}$ on X through

$$\mathcal{U} \in \mathcal{J}_{\mathcal{U}_X} \Leftrightarrow \mathcal{U}_X \subseteq \mathcal{U}. \tag{2.68}$$

However, not every uniform convergence structure is of the form (2.68). In this regard, we may recall [26] that a uniform convergence structure $\mathcal{J}_X$ on X induces a convergence structure $\lambda_{\mathcal{J}_X}$ on X, called the *induced convergence structure*, through

$$\begin{array}{ll} \forall & x \in X : \\ \forall & \mathcal{F} \text{ a filter on } X : \\ & \mathcal{F} \in \lambda_{\mathcal{J}_X}(x) \Leftrightarrow \mathcal{F} \times [x] \in \mathcal{J}_X \end{array}. \tag{2.69}$$

The induced convergence structure need not be a completely regular topology on X. In fact, every convergence structure λ_X which is *reciprocal*, that is,

$$\begin{array}{ll} \forall & x, y \in X : \\ & \lambda_X(x) = \lambda_X(y) \text{ or } \lambda_X(x) \cap \lambda_X(y) = \emptyset \end{array},$$

[5] In the original definition, this condition was replaced with the stronger one '$[\Delta] \in \mathcal{J}_X$'. This definition, however, results in a category which is not cartesian closed.

is induced by a uniform convergence structure through (2.69). Indeed, the family of filters $\mathcal{J}_{\lambda_X}$ on $X \times X$, called the *associated uniform convergence structure*, specified by

$$\mathcal{U} \in \mathcal{J}_{\lambda_X} \Leftrightarrow \left(\begin{array}{ll} \exists & x_1, ..., x_k \in X : \\ \exists & \mathcal{F}_1, ..., \mathcal{F}_k \text{ filters on } X : \\ & \quad 1) \quad \mathcal{F}_i \in \lambda_X (x_i) \text{ for } i = 1, ..., k \\ & \quad 2) \quad (\mathcal{F}_1 \times \mathcal{F}_1) \cap ... \cap (\mathcal{F}_k \times \mathcal{F}_k) \subseteq \mathcal{U} \end{array} \right) \tag{2.70}$$

is a complete uniform convergence structure that induces the convergence structure λ_X whenever it is reciprocal. In particular, every Hausdorff convergence structure is induced by the associated uniform convergence structure (2.70). This is clearly far more general than the completely regular topologies that are induced by uniformities.

The usual uniform concepts, namely, that of Cauchy filter, uniformly continuous function, completeness and completion extend in the natural way to uniform convergence spaces. In particular, if X and Y are uniform convergence spaces, then a mapping

$$f : X \to Y$$

is *uniformly continuous* whenever

$$\begin{array}{l} \forall \quad \mathcal{U} \in \mathcal{J}_X : \\ \quad (f \times f)(\mathcal{U}) \in \mathcal{J}_Y \end{array}$$

Furthermore, a uniformly continuous mapping is a *uniformly continuous embedding* if it is injective, and has a uniformly continuous inverse f^{-1} on the subspace $f(X)$ of Y, and a uniformly continuous embedding is a *uniformly continuous isomorphism* if it is also surjective.

The Cauchy sequences on a uniform convergence space are defined in the obvious way. A filter $\mathcal{F}$ on X is a Cauchy filter if and only if

$$\mathcal{F} \times \mathcal{F} \in \mathcal{J}_X, \tag{2.71}$$

and the uniform convergence space X is *complete* if every Cauchy sequence converges to some $x \in X$.

Weil's result on the completion of uniform spaces may be reproduced within the more general setting of uniform convergence spaces. In particular, Wyler [161] showed that with every Hausdorff uniform convergence space X one may associate a complete, Hausdorff uniform convergence space $X^\sharp$, and a uniformly continuous embedding

$$\iota_X : X \to X^\sharp$$

so that $\iota_X(X)$ is dense in $X^\sharp$. Furthermore, the completion $X^\sharp$ satisfies the *universal property* that, given any other complete, Hausdorff uniform convergence space Y, and a uniformly continuous mapping

$$f : X \to Y,$$

there is a uniformly continuous mapping

$$f^\sharp : X^\sharp \rightarrow Y^\sharp$$

so that the diagram

$$\begin{tikzcd} X \arrow[rr, "\iota_X"] \arrow[dr, "f"'] & & X^\sharp \arrow[dl, "f^\sharp"] \\ & Y & \end{tikzcd} \tag{2.72}$$

commutes. This completion is unique up to uniformly continuous isomorphism.

As we have mentioned, a uniformity on X is a particular case of a uniform convergence structure. In particular, with each such uniformity $\mathcal{U}_X$ on X we may associate a uniform convergence structure $\mathcal{J}_{\mathcal{U}_X}$ on X in a natural way through (2.68). As such, for each Hausdorff uniform space X, we may construct two completions, namely, the uniform space completion of Weil [160], and the uniform convergence space completion of Wyler [161]. These two completions, let us denote them by $X^\sharp_{We}$ and $X^\sharp_{Wy}$, respectively, are not the same. In particular, the Wyler completion $X^\sharp_{Wy}$ will typically not be a uniform space [161]. This apparent irregularity simply means that the Weil completion $X^\sharp_{We}$ will not satisfy the universal property enjoyed by the Wyler completion. More precisely, if X is a uniform space, and given a complete, Hausdorff uniform convergence space Y *which is not a uniform space*, and a uniformly continuous mapping

$$f : X \rightarrow Y$$

then we will in general not be able to find a uniformly continuous extension $f^\sharp : X^\sharp_{We} \rightarrow Y$ of f to $X^\sharp_{We}$.

Closely related to the concept of a uniform convergence space is that of a *Cauchy space*. Roughly speaking, a *Cauchy structure* on a set X is supposed to be the family of Cauchy filters associated with a given uniform convergence structure, and were introduced in an attempt to axiomatize the concept of Cauchy filter. These structures were axiomatized by Keller [80] as follows.

Definition 23 *Consider a set X, and a family $\mathcal{C}_X$ of filters on X. Then $\mathcal{C}_X$ is a Cauchy structure if it satisfies the following conditions:*

(i) $[x] \in \mathcal{C}_X$ *for each* $x \in X$.

(ii) *If* $\mathcal{F} \in \mathcal{C}_X$ *and* $\mathcal{F} \subseteq \mathcal{G}$, *then* $\mathcal{G} \in \mathcal{C}_X$.

(iii) *If* $\mathcal{F}, \mathcal{G} \in \mathcal{C}_X$ *and* $\mathcal{F} \vee \mathcal{G}$ *exists, then* $\mathcal{F} \cap \mathcal{G} \in \mathcal{C}_X$.

The pair $(X, \mathcal{C}_X)$ *is called a Cauchy space.*

Every uniform convergence structure $\mathcal{J}_X$ on a set X induced a unique Cauchy structure $\mathcal{C}_{\mathcal{J}_X}$ on X through

$$\mathcal{F} \in \mathcal{C}_{\mathcal{J}_X} \Leftrightarrow \mathcal{F} \times \mathcal{F} \in \mathcal{J}_X. \tag{2.73}$$

Conversely, every Cauchy structure $\mathcal{C}_X$ on X is induced by a uniform convergence structure. In particular, the family of filters $\mathcal{J}_{\mathcal{C}_X}$ on $X \times X$ defined as

$$\mathcal{U} \in \mathcal{J}_{\mathcal{C}_X} \Leftrightarrow \left(\begin{array}{ll} \exists & \mathcal{F}_1, ..., \mathcal{F}_k \in \mathcal{C}_X : \\ & (\mathcal{F}_1 \times \mathcal{F}_1) \cap ... \cap (\mathcal{F}_k \times \mathcal{F}_k) \end{array} \right) \tag{2.74}$$

constitutes a uniform convergence structure on X. Furthermore, the Cauchy structure induced by (2.74) is exactly $\mathcal{C}_X$.

It should be noted that, as is the case for uniform spaces, two different uniform convergence structures may induce the same Cauchy structures. In particular, the uniform convergence structure (2.74) is the largest uniform convergence structure, with respect to inclusion, that induces a given Cauchy structure. That is, if $\mathcal{J}_X$ is a uniform convergence structure that induces a given Cauchy structure through (2.69), then

$$\mathcal{J}_X \subseteq \mathcal{J}_{\mathcal{C}_X}. \tag{2.75}$$

It should be noted that several different 'completions' may be associated with a given Cauchy space, each with different properties [128]. Which completion is used is rather a matter of convenience. The Wyler completion is the unique completion that satisfies the universal extension property (2.72). However, the Wyler completion does not preserve compatibility with algebraic structures. In this regard, if X is a convergence vector space, then it carries in a natural way a uniform convergence structure. The underlying set associated with the Wyler completion $X^\sharp$ of X is a vector space in a straight forward way. However, in contradistinction with the Weil completion of a uniform space, the uniform convergence structure on the Wyler completion is not the one induced through by the algebraic structure on $X^\sharp$. This is also true for convergence groups [61]. Throughout the current work, we will always use the Wyler completion.

The role of uniform spaces, and more generally uniform convergence spaces, in analysis is well known. In particular, and most relevant to the current investigation, is the role played by such structures in the study of linear and nonlinear PDEs, as explained in Chapter 1, see also Chapter 6. Furthermore, these structures also appear in connection with the construction of compact spaces that contain a given topological space. In this regard, we may recall that Brummer and Hager [32] showed that, essentially, the Stone-Čech compactification of a completely regular topological space is in fact the completion of X equipped with a suitable uniformity.

Chapter 3

Real and Interval Functions

3.1 Semi-continuous Functions

The classical analysis of the nineteenth century was concerned mainly with sufficiently smooth functions, and in particular analytic functions. However, it is well known that even relatively simply constructions *involving only continuous functions* give rise to functions that are no longer continuous, see for instance [106] for an excellent historical overview of these and related matters. In this regard, consider the following example.

Example 24 *The pointwise limit of a sequence of continuous functions need not be continuous. Indeed, consider the sequence (u_n) of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ given by*

$$u_n(x) = \begin{cases} 0 & if \quad x \leq -\frac{1}{n} \\ \frac{nx+1}{2} & if \quad |x| < \frac{1}{n} \\ 1 & if \quad x \geq \frac{1}{n} \end{cases} . \tag{3.1}$$

Clearly the sequence (u_n) of continuous functions converges pointwise to the function

$$u(x) = \begin{cases} 0 & if \quad x < 0 \\ \frac{1}{2} & if \quad x = 0 \\ 1 & if \quad x > 0 \end{cases}$$

which has a discontinuity at $x = 0$.

Remark 25 *It should be noted that the pointwise limit of a sequence of continuous functions, such as that constructed in Example 24, may have discontinuities on a dense subset of the domain of convergence. In particular, in the case of real valued functions of a real variable, the limit will in general be continuous only on a residual set, which may have dense complement, see for instance [121] and [77].*

Furthermore, such discontinuous functions appear also in the applications of mathematics. In this regard, we may recall the discontinuous solution (1.12) to the nonlinear conservation law (1.4) to (1.5) which turns out to model the highly relevant physical phenomenon of shock waves.

An important class of such discontinuous functions arises in a natural way as a particular case of Example 24, namely, the semi-continuous real valued functions. The concept of a semi-continuous function generalizes that of a continuous function, and was first introduced by Baire [13] in the case of real valued functions of a real variable. Subsequently, the definition was extended to real valued functions on an arbitrary topological space, as well as functions with more general ranges, notably extended real valued functions, and set valued functions. In this case we will restrict our attention to the situation which is most relevant to the current investigation, namely, the case of extended real valued functions

$$u : X \rightarrow \overline{\mathbb{R}}$$

where X is a topological space, and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is the extended real line.

Recall that a function $u : X \rightarrow \mathbb{R}$ is continuous at $x \in X$ if and only if

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & V \in \mathcal{V}_x : \\ & y \in V \Rightarrow |u(x) - u(y)| < \epsilon \end{array} \tag{3.2}$$

which is equivalent to

$$\begin{array}{ll} \forall & \epsilon > 0 : \\ \exists & V \in \mathcal{V}_x : \\ & y \in V \Rightarrow u(x) - \epsilon < u(y) < u(x) + \epsilon \end{array} . \tag{3.3}$$

The concept of semi-continuity of a function $u : X \rightarrow \mathbb{R}$ at $x \in X$ is obtained by considering each of the two inequalities in (3.3) separately. In this regard, the standard definitions of lower semi-continuous function, and an upper semi-continuous function, respectively, are as follows.

Definition 26 *A function $u : X \rightarrow \overline{\mathbb{R}}$ is lower semi-continuous at $x \in X$ whenever*

$$\begin{array}{ll} \forall & M < u(x) : \\ \exists & V \in \mathcal{V}_x : \\ & y \in V \Rightarrow M < u(y) \end{array}$$

or $u(x) = -\infty$. If u is lower semi-continuous at every point of X, then it is lower semi-continuous on X.

Definition 27 *A function $u : X \rightarrow \overline{\mathbb{R}}$ is upper semi-continuous at $x \in X$ whenever*

$$\begin{array}{ll} \forall & M > u(x) : \\ \exists & V \in \mathcal{V}_x : \\ & y \in V \Rightarrow M > u(y) \end{array}$$

or $u(x) = +\infty$. If u is upper semi-continuous at every point of X, then it is upper semi-continuous on X.

Clearly, each continuous function $u : X \to \mathbb{R}$ is both lower semi-continuous and upper semi-continuous on X. Conversely, a function $u : X \to \mathbb{R}$ is both lower semi-continuous and upper semi-continuous on X, then it is continuous on X. Important examples of *discontinuous* semi-continuous functions include the indicator function χ_A of a set $A \subseteq X$, that is,

$$\chi_A(x) = \begin{cases} 1 & if \quad x \in A \\ & \\ 0 & if \quad x \notin A \end{cases} \tag{3.4}$$

If A is open, then χ_A is lower semi-continuous, and if A is closed, then χ_A is upper semi-continuous.

Semi-continuous functions appear as fundamental objects in analysis and its applications. In particular, such functions play a basic role in optimization theory, since semi-continuous functions have certain useful properties that fail in the case of continuous functions. In this regard, we may recall that the supremum of a set of continuous functions need not be continuous. In this regard, consider the following.

Example 28 *Consider the set $\{u_\alpha : \alpha > 1\}$ of continuous, real valued functions on $\mathbb{R}$, where each function u_α is defined by*

$$u_\alpha(x) = \begin{cases} 0 & if \quad |x| \geq \alpha \\ & \\ \frac{|x|-\alpha}{\alpha-1} & if \quad \alpha < |x| < 1 \\ & \\ -1 & if \quad |x| \leq 1 \end{cases}$$

The pointwise supremum of the set $\{u_\alpha : \alpha > 1\}$ is the function

$$u : \mathbb{R} \ni x \mapsto \sup\{u_\alpha(x) \ : \alpha > 1\} \in \mathbb{R}$$

which, in this case, is a well defined, real valued function given by

$$u_\alpha(x) = \begin{cases} 0 & if \quad |x| > 1 \\ & \\ -1 & if \quad |x| \leq 1 \end{cases}$$

which is not continuous on $\mathbb{R}$.[1]

Remark 29 *It should be noted that the function u constructed in Example 28 is continuous everywhere except on the closed nowhere dense set $\{\pm 1\} \subset \mathbb{R}$. In general, the pointwise supremum of a set of continuous functions may be discontinuous on a dense set. In particular, the set of points at which such a functions is continuous is in general only a residual set.*

[1]Similar examples may be constructed to show that the pointwise infimum of a set of continuous functions need not be continuous.

In contradistinction with the continuous case, semi-continuity of real valued functions is, to a certain extent, preserved when forming pointwise suprema and infima. In particular, if $\mathcal{A}$ is a set of lower semi-continuous functions on X, then the function $u : X \to \overline{\mathbb{R}}$ defined through

$$u : X \ni x \mapsto \sup\{v(x) \ : \ v \in \mathcal{A}\} \in \overline{\mathbb{R}} \tag{3.5}$$

is lower semi-continuous. Similarly, if $\mathcal{B}$ is a set of upper semi-continuous functions on X, then the function

$$l : X \ni x \mapsto \inf\{v(x) \ : \ v \in \mathcal{B}\} \in \overline{\mathbb{R}} \tag{3.6}$$

is upper semi-continuous. It should be noted that the infimum of a set of lower semi-continuous functions need not be lower semi-continuous, while the supremum of a set of upper semi-continuous functions is not always upper semi-continuous.

A particular case of (3.5) and (3.6) above occurs when the sets of functions $\mathcal{A}$ and $\mathcal{B}$ consist of continuous functions. Indeed, since a continuous function is both lower semi-continuous and upper semi-continuous, it follows immediately from (3.5) and (3.6) that

$$\begin{array}{l} \forall \quad \mathcal{A} \subset \mathcal{C}(X) \ : \\ \quad 1) \quad u : X \ni x \mapsto \sup\{v(x) \ : \ v \in \mathbb{A}\} \in \overline{\mathbb{R}} \text{ lower semi-continuous} \\ \quad 2) \quad l : X \ni x \mapsto \inf\{v(x) \ : \ v \in \mathbb{A}\} \in \overline{\mathbb{R}} \text{ upper semi-continuous} \end{array}$$

Conversely, if X is a metric space, then for each lower semi-continuous function $u : X \to \mathbb{R}$ we have

$$\begin{array}{l} \exists \quad \mathcal{A} \subset \mathcal{C}(X) \ : \\ \quad u(x) = \sup\{v(x) \ : \ v \in \mathcal{A}\}, \ x \in X \end{array},$$

while for every upper semi-continuous function $l : X \to \mathbb{R}$ we have

$$\begin{array}{l} \exists \quad \mathcal{B} \subset \mathcal{C}(X) \ : \\ \quad l(x) = \inf\{v(x) \ : \ v \in \mathcal{B}\}, \ x \in X \end{array}$$

Remark 30 *In general it is not true that the pointwise supremum, respectively infimum, of a set of continuous functions is the supremum, respectively infimum, of such a set with respect to the pointwise order on $\mathcal{C}(X)$. Indeed, if for each $n \in \mathbb{N}$ we define the function $u_n \in \mathcal{C}(\mathbb{R})$ through*

$$u_n(x) = \begin{cases} 1 - n|x| & if \quad |x| < \frac{1}{n} \\ \\ 0 & if \quad |x| \geq \frac{1}{n} \end{cases},$$

then the pointwise infimum of the set $\{u_n \ : \ n \in \mathbb{N}\}$ is the function

$$u(x) = \begin{cases} 1 & if \quad x = 0 \\ \\ 0 & if \quad x \neq 0 \end{cases},$$

while the infimum of the set $\{u_n \ : \ n \in \mathbb{N}\}$ in $\mathcal{C}(\mathbb{R})$ is the function which is identically 0 *on* $\mathbb{R}$.

As we have shown, semi-continuity of extended real valued functions is a far more general concept than continuity of such functions. However, such functions preserve certain useful properties of continuous functions. In particular, under certain mild assumptions on the function u, as well as the domain of definition X, the extreme value theorem for continuous functions may be generalized to semi-continuous functions. Indeed, if X is compact, and $u : X \to [-\infty, +\infty)$[2] is upper semi-continuous, then u attains a maximum value on X. That is,

$$\exists \quad \begin{array}{l} x_0 \in X : \\ u(x) \leq u(x_0),\ x \in X \end{array}.$$

Similarly, if $u : X \to (-\infty, +\infty]$ is lower semi-continuous, then u attains a minimum on X

$$\exists \quad \begin{array}{l} x_0 \in X : \\ u(x) \geq u(x_0),\ x \in X \end{array}.$$

Another useful property of continuous functions that extends to semi-continuous functions is concerned with the insertion of a continuous function in between two given functions. More precisely, given two continuous, real valued functions u and v on X such that $u \leq v$, then it is trivial observation that

$$\exists \quad \begin{array}{l} w \in \mathcal{C}(X) : \\ u \leq w \leq v \end{array}. \tag{3.7}$$

Indeed, we may simply take w to be the function $(u + v)/2$. In the nontrivial case when the continuous functions u and v in (3.7) are replaced with suitable semi-continuous functions, in particular u is upper semi-continuous and v is lower semi-continuous, (3.7) fails. However, a deep result due to Katětov [76] and Tong [152] characterizes normality of X in terms of such an insertion property. Namely, the topological space X is normal if and only if for each lower semi-continuous function v, and each upper semi-continuous function u such that $u \leq v$, we have

$$\exists \quad \begin{array}{l} w \in \mathcal{C}(X) : \\ u \leq w \leq v \end{array}. \tag{3.8}$$

This is a generalization of Hahn's Theorem [68], which states that (3.8) holds if X is a metric space.

Two fundamental operations associated with semi-continuous functions, and extended real valued functions in general, are the Baire Operators introduced by Baire [13] for real valued functions of a real variable. These operators were generalized to the case of extended real valued functions of a real variable by Sendov [146], and to functions defined on an arbitrary topological space by Anguelov [3]. Since these operators will be used extensively throughout the text, we include a detailed discussion of some of their more important properties.

[2] Note that, in case the function is allows to assume the value $+\infty$ on X, the result is trivial.

In this regard, we denote by $\mathcal{A}(X)$ the set of extended real valued functions on a topological space X. That is,

$$\mathcal{A}(X) = \{u : X \to \overline{\mathbb{R}}\}.$$

The Lower Baire Operator I and Upper Baire Operator S are, respectively, mappings

$$I : \mathcal{A}(X) \to \mathcal{A}(X)$$

and

$$S : \mathcal{A}(X) \to \mathcal{A}(X)$$

which are defined as

$$I(u)(x) = \sup\{\inf\{u(y) \ : \ y \in V\} : V \in \mathcal{V}_x\},\ x \in X \tag{3.9}$$

and

$$S(u)(x) = \inf\{\sup\{u(y) \ : \ y \in V\} : V \in \mathcal{V}_x\},\ x \in X, \tag{3.10}$$

respectively, with $\mathcal{V}_x$ denoting the neighborhood filter at $x \in X$. The connection of the operators (3.9) and (3.10) with semi-continuity of functions in $\mathcal{A}(X)$ may be seen immediately. Indeed, the Baire Operators *characterize* semi-continuity through

$$u \in \mathcal{A}(X) \text{ is lower semi-continuous } \Leftrightarrow I(u) = u \tag{3.11}$$

and

$$u \in \mathcal{A}(X) \text{ is upper semi-continuous } \Leftrightarrow S(u) = u, \tag{3.12}$$

respectively. Furthermore, for each $u \in \mathcal{A}(X)$ the function $I(u)$ is lower semi-continuous, while the function $S(u)$ is upper semi-continuous. From (3.11) and (3.12) it therefore follows that the operators I and S are *idempotent*, that is, for each $u \in \mathcal{A}(X)$

$$I(I(u)) = I(u) \tag{3.13}$$

and

$$S(S(u)) = S(u). \tag{3.14}$$

Moreover, from the definitions (3.9) and (3.10) it is clear that the operators I and S are also *monotone* with respect to the pointwise order on $\mathcal{A}(X)$

$$\begin{array}{l} \forall \quad u, v \in \mathcal{A}(X) : \\ \quad u \leq v \Rightarrow \left(\begin{array}{ll} 1) & I(u) \leq I(v) \\ 2) & S(u) \leq S(v) \end{array} \right) \end{array}. \tag{3.15}$$

It follows from (3.15) that compositions of the operators I and S must also be monotone so that

$$\begin{array}{l} \forall \quad u, v \in \mathcal{A}(X) \ : \\ u \leq v \Rightarrow \left(\begin{array}{ll} 1) & (I \circ S)(u) \leq (I \circ S)(v) \\ 2) & (S \circ I)(u) \leq (S \circ I)(v) \end{array} \right). \end{array} \tag{3.16}$$

Furthermore, the composite operators in (3.16) are also idempotent so that

$$(I \circ S)((I \circ S)(u)) = (I \circ S)(u) \tag{3.17}$$

and

$$(S \circ I)((S \circ I)(u)) = (I \circ S)(u) \tag{3.18}$$

for each $u \in \mathcal{A}(X)$. Indeed, by the obvious inequality

$$I(u) \leq S(u), \tag{3.19}$$

as well as (3.13) to (3.14), we have

$$(I \circ S) \circ (I \circ S)(u) \leq (I \circ S) \circ (S \circ S)(u) = (I \circ S)(u)$$

and

$$(I \circ S) \circ (I \circ S)(u) \geq (I \circ I) \circ (I \circ S)(u) = (I \circ S)(u)$$

which implies (3.17). The identity (3.18) is obtained by similar arguments.

A particularly useful class of semi-continuous functions is that of the normal semi-continuous functions. These functions were introduced by Dilworth [47] in connection with his attempts at obtaining a representation of the Dedekind order completion of spaces of continuous functions. In particular, Dilworth showed that the Dedekind order completion of the space $\mathcal{C}_b(X)$ of all *bounded*, real valued continuous functions on a completely regular topological space X is the set of bounded normal upper semi-continuous functions on X.

A definition of normal semi-continuity for arbitrary real valued functions was given by Anguelov [3], which coincides with Dilworth's definition in the case of bounded functions. This is the definition that we will use, and it is most simply stated in terms of the Baire Operators I and S. Namely, a real valued function $u \in \mathcal{A}(X)$ is *normal lower semi-continuous* whenever

$$(I \circ S)(u) = u \tag{3.20}$$

and it is *normal upper semi-continuous* whenever

$$(I \circ S)(u) = u \tag{3.21}$$

From the time Dilworth introduced normal semi-continuous functions in the 1950s, up to very recently, there was rather limited interest in these functions. There are two possible reasons for such a lack of enthusiasm concerning normal semi-continuous functions functions. Firstly, Dilworth's results on the order completion of spaces of continuous functions applies only in rather particular cases, namely, when the underlying space X is compact, or to the set of *bounded* continuous functions on a completely regular space. Moreover, it appeared to be rather difficult to extend Dilworth's results to the case of unbounded functions. Several attempts at more general results concerning the order completion of spaces of continuous functions have resulted only in partial success, see for instance [56]. Furthermore, these functions appeared at the time not to have other significant applications.

Recently, as will be discussed in more detail in Section 3.2, there has been a renewed interest in such functions. The current interest in these functions is due to two highly nontrivial applications of normal semi-continuous functions. Indeed, Anguelov [3] significantly extended Dilworth's results on the Dedekind order completion of spaces of continuous functions, using spaces of functions that are essentially equivalent to normal semi-continuous functions. This, in turn, has lead to a significant improvement of the regularity of generalized solutions to large classes of nonlinear PDEs obtained through the Order Completion Method [8], [9].

3.2 Interval Valued Functions

The field of interval analysis, and in particular interval valued functions, is a subject that is traditionally associated with validated computing [2], [87], [146]. The central issue is to design algorithms generating bounds for exact solutions of mathematical problems, and such bounds may be represented as intervals, and interval valued functions. In this context, such functions appear in a natural way as error bounds for numerical and theoretical computations. Such interval valued functions have also been applied to approximation theory [146]. In fact, Sendov [146] introduced the concept of a Hausdorff continuous interval valued function in connection with Hausdorff approximations of real functions of a real argument. However, recent applications of interval valued functions to diverse mathematical fields previously considered to be unrelated to interval analysis, have lead to a renewed interest in these functions, as well as to a new point of view regarding them. Namely, the possible structures, of whatever appropriate kind (topological, algebraic or order theoretic), with which *spaces* of interval valued functions may be equipped.

Let us now briefly recall the basic notations and concepts involved. In this regard, we denote by

$$\overline{\mathbb{IR}} = \left\{ a = [\underline{a}, \overline{a}] \;\middle|\; \begin{array}{ll} 1) & \underline{a}, \overline{a} \in \overline{\mathbb{R}} \\ 2) & \underline{a} \leq \overline{a} \end{array} \right\}$$

the set of extended, closed real intervals. The subset of $\overline{\mathbb{IR}}$ consisting of finite and closed real intervals is denoted by $\mathbb{IR}$. By identifying a point $a \in \overline{\mathbb{R}}$ with the degenerate interval $[a,a] \in \overline{\mathbb{IR}}$, we may consider the extended real line as a subset of $\overline{\mathbb{IR}}$

$$\overline{\mathbb{R}} \subset \overline{\mathbb{IR}}. \tag{3.22}$$

The usual *total order* on $\overline{\mathbb{R}}$ can be extended to a *partial order* on $\overline{\mathbb{IR}}$ in several different ways. A particularly useful order was defined by Markov [104] though

$$a \le b \Leftrightarrow \left(\begin{array}{ll} 1) & \underline{a} \le \underline{b} \\ 2) & \overline{a} \le \overline{b} \end{array} \right) \tag{3.23}$$

Our main interest in this section is in functions whose values are extended real intervals. For a given set X we denote by $\mathbb{A}(X)$ the set of such interval valued functions on X, that is,

$$\mathbb{A}(X) = \{u : X \to \overline{\mathbb{IR}}\}. \tag{3.24}$$

A convenient representation of interval valued functions is as pairs of point valued functions. In particular, with every $u \in \mathbb{A}(X)$ we associate the pair of point valued functions $\underline{u}, \overline{u} \in \mathcal{A}(X)$ such that

$$u(x) = [\underline{u}(x), \overline{u}(x)],\ x \in X. \tag{3.25}$$

Through the identification of points in $\overline{\mathbb{R}}$ with intervals in $\overline{\mathbb{IR}}$, we may consider the set $\mathcal{A}(X)$ of extended real valued functions on X as a subset of $\mathbb{A}(X)$. Since the partial order (3.23) on $\overline{\mathbb{IR}}$ extends the usual total order on $\overline{\mathbb{R}}$, the pointwise order on $\mathbb{A}(X)$, specified as

$$u \le v \Leftrightarrow \left(\begin{array}{ll} \forall & x \in X : \\ & u(x) \le v(x) \end{array} \right), \tag{3.26}$$

extends the pointwise order on $\mathcal{A}(X)$.

Several concepts of *continuity* of interval valued functions, defined on a topological space X, have been introduced in the literature [6], [146]. Here we may recall the concepts of Hausdorff continuity, Dilworth continuity and Sendov continuity, all of which are closely linked to the concepts of semi-continuity of extended real valued functions discussed in Section 3.1. These continuity concepts are conveniently formulated in terms of extensions of the Baire operators (3.9) and (3.10). In this regard, we note that these operators act in a natural way also on interval valued functions. Indeed, for $u \in \mathbb{A}(X)$ we may define the operators I and S as

$$I(u)(x) = \sup\{\inf\{\underline{u}(y) \ : \ y \in V\} : V \in \mathcal{V}_x\},\ x \in X \tag{3.27}$$

and

$$S(u)(x) = \inf\{\sup\{\overline{u}(y) \ : \ y \in V\} : V \in \mathcal{V}_x\},\ x \in X \tag{3.28}$$

which clearly coincides with (3.9) and (3.10), respectively, if u is point valued.

It should be noted that the extended Baire operators (3.27) and (3.28) produce point valued functions. That is,

$$I : \mathbb{A}(X) \to \mathcal{A}(X)$$

and

$$S : \mathbb{A}(X) \to \mathcal{A}(X)$$

As such, and in view of (3.19), the Graph Completion Operator

$$F : \mathbb{A}(X) \ni u \mapsto [I(u), S(u)] \in \mathbb{A}(X), \tag{3.29}$$

introduced by Sendov [146] for finite interval valued functions of a real argument, see also [3], is a well defined mapping. From (3.15), (3.13) and (3.14) it follows that the operator F is both *monotone* and *idempotent*, that is, for all $u \in \mathbb{A}(X)$

$$F(F(u)) = F(u), \tag{3.30}$$

and for all $u, v \in \mathbb{A}(X)$

$$u \leq v \Rightarrow F(u) \leq F(v). \tag{3.31}$$

An important class of interval valued functions, namely, the Sendov continuous (S-continuous) functions, is defined as the fixed points of the operator F. That is, $u \in \mathbb{A}(X)$ is S-continuous if and only if

$$F(u) = u. \tag{3.32}$$

These functions, introduced by Sendov, play an important role in the theory of Hausdorff approximations [146].

The class of functions of main interest in the current context is that of the Hausdorff continuous functions, which are defined as follows [146], see also [3].

Definition 31 *A function $u \in \mathbb{A}(X)$ is Hausdorff continuous (H-continuous) if*

$$F(u) = u \tag{3.33}$$

and for each $v \in \mathbb{A}(X)$

$$\left(\begin{array}{ll} \forall & x \in X : \\ & v(x) \subseteq u(x) \end{array} \right) \Rightarrow F(v) = u. \tag{3.34}$$

The set of all H-continuous functions on X is denoted as $\mathbb{H}(X)$.

Remark 32 *A remark on the particular meanings of conditions (3.33) and (3.34) is appropriate. Essentially, the condition (3.33) may be interpreted as a continuity requirement. Indeed, in view of the definition (3.29) of the Graph Completion Operator and the characterization (3.12) of semi-continuity in terms of the Baire operators, the condition (3.33) simply states that the function u may be represented by a pair of semi-continuous functions, namely, $\underline{u}$ is lower semi-continuous, and $\overline{u}$ is upper semi-continuous. The second condition (3.34) is in fact a minimality condition with respect to inclusion. That is, u is smallest S-continuous function, with respect to inclusion, included in u.*

H-continuous functions were first introduced by Sendov [146] for functions of a real variable in connection with applications to the theory of Hausdorff approximations. Recently [3], the definition was extended to functions defined on arbitrary topological spaces.

The set $\mathbb{H}(X)$ of H-continuous functions inherits the partial order (3.26). With this order, the set $\mathbb{H}(X)$ is a *complete lattice*, that is,

$$\begin{array}{ll} \forall & A \subseteq \mathbb{H}(X) \ : \\ \exists & u_0, l_0 \in \mathbb{H}(X) \ : \\ & \quad 1) \quad u_0 = \sup A \\ & \quad 2) \quad l_0 = \inf A \end{array} . \tag{3.35}$$

Furthermore, the supremum and infimum in (3.35) may be described in terms of the pointwise supremum and infimum

$$\varphi : X \ni x \mapsto \sup\{u(x) \ : \ u \in A\} \in \overline{\mathbb{R}}$$

and

$$\psi : X \ni x \mapsto \inf\{u(x) \ : \ u \in A\} \in \overline{\mathbb{R}},$$

respectively. Indeed, see [3], we have the following characterization of u_0 and l_0 in (3.35):

$$u_0 = F(I(S(\varphi))), \, l_0 = F(S(I(\psi)))$$

To date, three important classes of H-continuous functions have been identified. These are the *finite* H-continuous functions $\mathbb{H}_{ft}(X)$, the *bounded* H-continuous functions $\mathbb{H}_b(X)$ and the *nearly finite* H-continuous functions $\mathbb{H}_{nf}(X)$. These classes of functions are defined as

$$\mathbb{H}_{ft}(X) = \left\{ u \in \mathbb{H}(X) \left| \begin{array}{ll} \forall & x \in X : \\ & u(x) \in \mathbb{I}\mathbb{R} \end{array} \right. \right\}, \tag{3.36}$$

$$\mathbb{H}_b(X) = \left\{ u \in \mathbb{H}_{ft}(X) \left| \begin{array}{ll} \exists & [\underline{a}, \overline{a}] \in \mathbb{I}\mathbb{R} : \\ & u(x) \subseteq [\underline{a}, \overline{a}], \, x \in X \end{array} \right. \right\} \tag{3.37}$$

and

$$\mathbb{H}_{nf}(X) = \left\{ u \in \mathbb{H}(X) \middle| \begin{array}{l} \exists \quad \Gamma \subset X \text{ closed nowhere dense :} \\ \quad x \in X \setminus \Gamma \Rightarrow u(x) \in \mathbb{IR} \end{array} \right\}, \tag{3.38}$$

respectively. The relevance of these classes of functions is evident from the recent and highly nontrivial applications to diverse branches of mathematics [3], [8], [9], [10]. One of the applications, namely, the the Order Completion Method for nonlinear PDEs [8], [9] is recounted in Section 1.4, while the discussion of an application to topological completion of $\mathcal{C}(X)$ [10] is postponed to Chapter 4. Here we proceed with a short account of the main results of Anguelov [3] concerning the Dedekind order completion of $\mathcal{C}(X)$.

In this regard, we note that, as a simple corollary to (3.35), each of the spaces (3.36) to (3.38) is a Dedekind order complete lattice (1.96) with respect to the order induced from $\mathbb{H}(X)$. Furthermore, each *continuous*, real valued function on X is also H-continuous. Indeed, since each continuous, real valued function u is both lower semi-continuous and upper semi-continuous, it follows by (3.11) and (3.12) as well as the definition (3.29) of the Graph Completion Operator that $F(u) = u$. Furthermore, since u is point valued, it follows that the second condition (3.34) of Definition 31 also holds. Then the set of inclusions

$$\mathcal{C}_b(X) \subseteq \mathbb{H}_b(X), \mathcal{C}(X) \subseteq \mathbb{H}_{ft}(X) \tag{3.39}$$

is obvious. Moreover, in view of the fact that the order (3.26) extends the usual pointwise order on $\mathcal{A}(X)$, it is clear that the inclusions in (3.39) are in fact also order isomorphic embeddings (51). What remains to be verified is the denseness properties

$$\begin{array}{ll} \forall & u \in \mathbb{H}_{ft}(X) : \\ & u = \sup \{ v \in \mathcal{C}(X) \ : \ v \leq u \} \end{array} \tag{3.40}$$

and

$$\begin{array}{ll} \forall & u \in \mathbb{H}_b(X) : \\ & u = \sup \{ v \in \mathcal{C}_b(X) \ : \ v \leq u \} \end{array}. \tag{3.41}$$

The order denseness properties (3.41) holds for every topological space X. The property (3.40) is valid whenever X is a metric space, or when X is completely regular and satisfies

$$\begin{array}{ll} \forall & u \in \mathbb{H}_{nf}(X) : \\ \exists & v \in \mathcal{C}(X) : \\ & u \leq v \end{array}.$$

In particular, in each of these cases, we may construct for each $u \in \mathbb{H}_{ft}(X)$ sequences (λ_n) and (μ_n) of continuous functions on X so that

$$\begin{array}{ll} \forall & n \in \mathbb{N} : \\ & \lambda_n \leq \lambda_{n+1} \leq u \leq \mu_{n+1} \leq \mu_n \end{array} \tag{3.42}$$

and

$$\begin{array}{ll} \forall & x \in X : \\ & 1) \quad \sup\{\lambda_n(x) \,:\, n \in \mathbb{N}\} = \underline{u}(x) \\ & 2) \quad \inf\{\mu_n(x) \,:\, n \in \mathbb{N}\} = \overline{u}(x) \end{array} \tag{3.43}$$

Now, as mentioned, the H-continuous functions are essentially equivalent to suitable normal semi-continuous functions. In this regard, consider a finite H-continuous function u on X. Then clearly, the function

$$\underline{u} = (I \circ S)(u) \tag{3.44}$$

is real valued and normal lower semi-continuous. Conversely, for a given real valued normal lower semi-continuous function $\underline{v}$ on X, the function

$$v = F(\underline{v})$$

is a finite H-continuous function. Indeed, the mapping

$$F : \mathcal{NL}_{ft}(X) \to \mathbb{H}_{ft}(X) \tag{3.45}$$

is a bijection, with $\mathcal{NL}_{ft}(X)$ the space of all real valued normal lower semi-continuous functions on X. Moreover, given normal lower semi-continuous functions $\underline{u}$ and $\underline{v}$ on X, we have

$$\underline{u} \leq \underline{v} \Leftrightarrow F(\underline{u}) \leq F(\underline{v}) \tag{3.46}$$

so that F is in fact an order isomorphism.

Remark 33 *The same construction may be reproduced for normal upper semi-continuous functions. In this case, one may consider the mapping*

$$F : \mathcal{NU}_{ft}(X) \to \mathbb{H}_{ft}(X)$$

instead of (3.45), where $\mathcal{NU}_{ft}(X)$ denotes the set of real valued normal upper semi-continuous functions.

The concept of an interval valued function, and in particular that of an H-continuous function, proves to be a useful concept in so far as the representations of certain extensions of spaces of continuous functions are concerned. Here we discussed only the extension through order, but recently the rational completion of the ring $\mathcal{C}(X)$ and other related extensions of this space were also characterized as spaces of H-continuous functions [4]. Moreover, and as will be discussed in more detail in Chapter 4, the convergence vector space completion of $\mathcal{C}(X)$ with respect to the so called order convergence structure may be constructed as the set $\mathbb{H}_{ft}(X)$. However, in view of the fact that interval valued functions are relatively unknown in analysis, our preference in the current investigation is rather to use the equivalent representation via normal lower semi-continuous functions.

Chapter 4

Order and Topology

4.1 Order, Algebra and Topology

Order, together with algebra and topology, are the most fundamental concepts in modern mathematics, and the rich mathematical field that may be broadly called 'analysis' is an example of the great power and utility of mathematical concepts that arise as combinations of these basic notions. In this regard, we may recall the theory of Operator Algebras, initiated by von Neumann, where all three these basic concepts are involved. In this section, we will discuss two examples of such fruitful interactions between the basic trio of order, algebra and topology which is relevant to the current investigation. Namely, the theory of ordered algebraic structures, in particular Riesz Spaces, and that of ordered topological spaces.

Many of the interesting spaces in analysis, and in particular linear functional analysis, are equipped with partial orders in a natural way. Some of the most prominent examples include the following. The space $\mathcal{C}(X)$ of continuous, real valued functions on a topological space X ordered in the usual pointwise way

$$\begin{array}{l} \forall \quad u, v \in \mathcal{C}(X) \; : \\ \quad u \leq u \Leftrightarrow \left(\begin{array}{ll} \forall & x \in X : \\ & u(x) \leq v(x) \end{array} \right) , \end{array}$$

and the space $\mathbf{M}(\Omega)$ of real valued measurable functions on a measure space (Ω, Λ, μ), modulo functions that are almost everywhere equal, equipped with the almost everywhere pointwise order

$$\begin{array}{l} \forall \quad u, v \in \mathbf{M}(\Omega) \\ \quad u \leq v \Leftrightarrow \left(\begin{array}{ll} \exists & E \subset \Omega, \; \mu(E) = 0 : \\ & u(x) \leq v(x), \; x \in \Omega \setminus E \end{array} \right) \end{array}$$

as well as its important subspaces of p-integrable functions. These are defined as

$$L_p(\Omega) = \left\{ u \in \mathbf{M}(\Omega) \; : \; \int_\Omega |u(x)|^p dx < \infty \right\}$$

for $p \leq 1 < \infty$, and

$$L^{\infty}(\Omega) = \left\{ u \in \mathbf{M}(\Omega) \left| \begin{array}{ll} \exists & C > 0 : \\ \exists & E \subset \Omega,\ \mu(E) = 0 : \\ & u(x) \leq C,\ x \in \Omega \setminus E \end{array} \right. \right\}$$

The spaces described above are all examples of Riesz spaces, also called vector lattices. A Riesz space is a real vector space L equipped with a partial order in such a way that L is a lattice and

$$\begin{array}{ll} \forall & u, v, w \in L : \\ \forall & \alpha \in \mathbb{R},\ \alpha \geq 0 : \\ & 1) \quad u \leq v \Rightarrow u + w \leq v + w \\ & 2) \quad u \geq 0 \Rightarrow \alpha u \geq 0 \end{array} . \tag{4.1}$$

Riesz spaces were introduced independently, and more or less simultaneously, by F. Riesz [130], [131] , L. V. Kantorovitch [72], [73], [74] and H. Freudenthal [60].

The simple requirements on the compatibility of the order on L and its algebraic structure in (4.1) has some immediate and rather unexpected consequences. In this regard, we may recall [101] that any Riesz space L is a fully distributive lattice. That is,

$$\begin{array}{ll} \forall & A \subset L : \\ \forall & u \in L : \\ & \inf\{\sup\{u, v\} : v \in A\} = \inf\{u, \sup A\} \end{array}$$

provided that $\sup A$ exists in L.

The utility and power of methods from Riesz space theory, when applied to problems in analysis and other parts of mathematics, is well documented, although often not fully appreciated. In this regard, we may recall Freudenthal's Spectral Theorem [60], see also [101]. Roughly speaking, Freudenthal's Theorem states that each $u \in L$ may be approximated, in a suitable sense, by 'step functions'. The importance of this result is clear from its applications. These include the following three fundamental results, namely, the Radon-Nykodym Theorem in measure theory, the Spectral Theorem for Hermitian operators and normal operators in Hilbert space, and the Poisson formula for harmonic functions on an open circle, see [101].

The theory of Riesz spaces has been extensively developed, and most of the basic questions have by now been settled for nearly twenty years. The most recent theoretical program was initiated by A. C. Zaanen in the later part of the 1980s. The aim of this program was to reprove the major results of the theory in elementary terms, without reference to the often highly complicated representation theorems for Riesz spaces. This area of research has culminated in 1997 in the excellent introductory text [164]. Current interest in Riesz spaces stem from their many applications, in particular in connection with stochastic calculus and martingale theory, see for instance [45], [91], [92] and [95].

A useful generalization of the concept of a Riesz space is that of a lattice ordered group, called an l-group for short. In this regard, recall [29] that an l-group is a commutative group $G = (G, +,)$ equipped with a lattice order $\leq$ such that

$$\begin{array}{ll} \forall & f, g, h \in G : \\ & f \leq g \Rightarrow \left(\begin{array}{ll} 1) & f + h \leq g + h \\ 2) & h + f \leq h + g \end{array} \right) \end{array} . \tag{4.2}$$

Such an l-group is called Archimedean whenever

$$\begin{array}{ll} \forall & f, g \geq 0 : \\ & nf \leq g,\ n \in \mathbb{N} \Rightarrow f = 0 \end{array} , \tag{4.3}$$

with 0 the group identity in G. Loosely speaking, the condition (4.3) means that the group G does not contain any 'infinitely small' of 'infinitely large' elements. Lastly, an element $e \in G$ is an order unit if $e > 0$ and

$$\begin{array}{ll} \forall & g \in G\ ,\ g \geq 0 : \\ \exists & n \in \mathbb{N} : \\ & g \leq ne \end{array} ,$$

while e is a weak order unit whenever $e > 0$ and

$$\begin{array}{ll} \forall & g \in G,\ g \geq 0 : \\ & \inf\{e, g\} = 0 \Rightarrow g = 0 \end{array} .$$

A useful aspect of the theory of Archimedean l-groups is that, each such group admits a representation as a set of continuous, real valued functions on a suitable topological space. In particular, if G admits a distinguished weak order unit, then we may associate with G a completely regular topological space, call it X_G, and an l-group homomorphism

$$T_G : G \to \mathcal{C}(X_G) .$$

That is, every commutative l-group with distinguished weak order unit is isomorphic to an l-subgroup of the space of continuous functions on certain completely regular topological space X_G, see [162] as well as [16] and the references cited there. Conversely, we may associate with each completely regular topological space X a commutative l-group with weak order unit G_X, namely, the l-group $\mathcal{C}(X)$. This amounts to a relationship between the categories W of Archimedean l-groups with weak order unit, and the category R of completely regular topological spaces. As such, one may study the latter category through the more simple one W.

On a partially ordered set X there are several ways to define a topology in terms of the order on X, see for instance [29]. Among these, we may mention the order topology [82], the interval topology [83], the Scott topology and the Lawson topology [66]. Such topologies turn out to be interesting from the point of view of applications. In this regard, we may recall that the Scott topology and the

Lawson topology play an important role in the theory of continuous lattices, and its applications to theoretical computer science [66].

More generally, given a topological space X equipped with a partial order, we may require the order to be compatible with the topology in a suitable sense. One particularly useful requirement is that the mappings

$$\vee : X \times X \ni (x, y) \mapsto \sup\{x, y\} \in X \tag{4.4}$$

and

$$\wedge : X \times X \ni (x, y) \mapsto \inf\{x, y\} \in X \tag{4.5}$$

be continuous on their respective domains of definition, with respect to the product topology on $X \times X$. Such a space is termed an ordered topological space. The order topology, and the interval topology on a lattice X, see for instance [53], are both examples of ordered topological spaces. It turns out that, in this case, many properties of the topology τ on X may be characterized in terms of the partial order. This may represent a dramatic simplification in so far as topological concepts may be described at the significantly more simple level of order.

4.2 Convergence on Posets

As mentioned in Chapter 2, we may associate with every topology τ on a set X a notion of convergence with respect to τ. However, one may also define several useful and important concepts of convergence that cannot be associated with a topology. In particular, we may, for example, associate with each element x of a set X a collection $\sigma(x)$ of sequences on X, which is interpreted to mean that a sequence (x_n) converges to x if and only if (x_n) belongs to $\sigma(x)$. Such an association of sequences with points is in general not determined by a topology.

In the previous section, we discussed the idea of defining a topology in terms of a partial order. In this case, and in view of the above remark, we may associate with the partial order the convergence induced by the topology. More generally, we may define a notion of convergence of sequences in terms of a given partial order on a set X. Indeed, several such useful notions of convergence on partially ordered sets have been introduced in the literature, see for instance [29], [50], [101] and [105]. It often happens that these notions of convergence cannot be associated with any topology, see for instance [154].

One of the most well known such examples of convergence defined through a partial order that is, in general, not associated with any topology, is order convergence of sequences, see for instance [101] or [105]. In this regard, see also Example 9, for a given partially ordered set X, a sequence (x_n) order converges to $x \in X$ whenever

$$\begin{array}{ll} \exists & (\lambda_n),\ (\mu_n) \subset X : \\ & 1) \quad n \in \mathbb{N} \Rightarrow \lambda_n \leq \lambda_{n+1} \leq x_{n+1} \leq \mu_{n+1} \leq \mu_n \\ & 2) \quad \sup\{\lambda_n : n \in \mathbb{N}\} = x = \inf\{\mu_n : n \in \mathbb{N}\} \end{array} \tag{4.6}$$

In general, there is no topology τ on X such that, for each $x \in X$, and every sequence (x_n) on X, (x_n) order converges to x if and only if (x_n) converges to x with respect to τ, see for instance [154]. Indeed, the following two properties of convergent sequences in topological spaces [107] fail to hold for the order convergence of sequences. Namely, the Divergence Axiom

> If every subsequence of (x_n) contains a subsequence which converges to $x \in X$, then (x_n) converges to x.

and the Axiom of Iterated Limits

> If, for every $n \in \mathbb{N}$, the sequence $(x_{n,m})$ converges to $x_n \in X$, and the sequence (x_n) converges to x in X, then there is a strictly increasing sequence of natural numbers (m_n) so that the sequence (x_{n,m_n}) converges to x in X. (4.7)

In this regard, we may recall Example 9. Note, however, that the following version of the Axiom of Iterated Limits (4.7) remains valid under rather general conditions on the partially ordered set X.

Proposition 34 **[10] Let L be a lattice with respect to a given partial order $\leq$.*

1. *For every $n \in \mathbb{N}$, let the sequence $(u_{m,n})$ in L be bounded and increasing and let*

$$u_n = \sup\{u_{m,n} : m \in \mathbb{N}\},\ n \in \mathbb{N}$$
$$u'_n = \sup\{u_{m,n} : m = 1, ..., n\}$$

If the sequence (u_n) is bounded from above and increasing, and has supremum in L, then the sequence (u'_n) is bounded and increasing and

$$\sup\{u_n : n \in \mathbb{N}\} = \sup\{u'_n : n \in \mathbb{N}\}$$

2. *For every $n \in \mathbb{N}$, let the sequence $(v_{m,n})$ in L be bounded and decreasing and let*

$$v_n = \inf\{v_{m,n} : m \in \mathbb{N}\},\ n \in \mathbb{N}$$
$$v'_n = \inf\{v_{m,n} : m = 1, ..., n\}$$

If the sequence (v_n) is bounded from below and decreasing, and has infimum in L, then the sequence (v'_n) is bounded and decreasing and

$$\inf\{v_n : n \in \mathbb{N}\} = \inf\{v'_n : n \in \mathbb{N}\}$$

Since the usual Hausdorff concept of topology is insufficient to describe the order convergence of sequences, the question arises whether or not convergence structures provide a sufficiently general context for the study of the order convergence of sequences. Several authors have addressed this issue, and similar problems arising in connections with other types of convergence on partially ordered sets, see for example [55] and [67]. In this regard, R Ball [15] showed that the order convergence of (generalized) sequences, on an l-group is induced by a group convergence structure. In particular, it is shown that the convergence group completion of such an l-group with respect to the mentioned group convergence structure is the Dedekind order completion of G. Papangelou [122] considered a similar problem in the setting of sequential convergence groups.

Recently [10], [155], it was shown that for any σ-distributive lattice X, that is, a lattice X that satisfies

$$\begin{array}{ll} \forall & (x_n) \subseteq X : \\ \forall & x \in X : \\ & \sup\{x_n : n \in \mathbb{N}\} = x_0 \Rightarrow \sup\{\inf\{x, x_n\} : n \in \mathbb{N}\} = \inf\{x, x_0\} \end{array},$$

there is a convergence structure on X that induces the order convergence of sequences. In particular, the convergence structure λ_o, specified as

$$\begin{array}{ll} \forall & x \in X : \\ \forall & \mathcal{F} \text{ a filter on } X : \\ & \mathcal{F} \in \lambda_o(x) \Leftrightarrow \left(\begin{array}{ll} \exists & (\lambda_n), (\mu_n) \subset X : \\ & 1) \quad n \in \mathbb{N} \Rightarrow \lambda_n \leq \lambda_{n+1} \leq x \leq \mu_{n+1} \leq \mu_n \\ & 2) \quad \sup\{\lambda_n : n \in \mathbb{N}\} = x = \inf\{\mu_n : n \in \mathbb{N}\} \\ & 3) \quad [\{[\lambda_n, \mu_n] : n \in \mathbb{N}\}] \subseteq \mathcal{F} \end{array} \right) \end{array} \quad (4.8)$$

is first countable, Hausdorff, and induces the order convergence of sequences. In particular, if X is an Archimedean Riesz space, then the convergence structure (4.8) is a vector space convergence structure, and X a convergence vector space. In this case, we may construct the convergence vector space completion of X, which is the Dedekind σ-completion of X [155], equipped with the order convergence structure (4.8). In particular, in case X is the set $\mathcal{C}(Z)$ of continuous functions on a metric space Z, then the completion may be constructed as the set $\mathbb{H}_{ft}(Z)$ of finite H-continuous interval valued functions on Z.

Chapter 5

Organization of the Thesis

5.1 Objectives of the Thesis

The aim of this thesis is to develop a general and type independent theory concerning the existence and regularity of generalized solutions of systems of nonlinear PDEs. In this regard, our point of departure is the Order Completion Method [119], which is discussed in Section 1.4. In particular, this includes, as a first and basic step, a reformulation of the Order Completion Method in the context of uniform convergence spaces. That is, the construction of generalized solutions to a nonlinear PDE (1.100) as an element of the Dedekind completion of the space $\mathcal{M}^m_{\mathbf{T}}(\Omega)$ is interpreted in terms of the Wyler completion of a suitable uniform convergence space. Such a recasting of the Order Completion Method in terms of uniform convergence spaces allows for the application of convergence theoretic techniques to problems relating to the structure and regularity of generalized solutions, techniques that may turn out to be more suitable to the mentioned problems that the order theoretic methods involved in the Order Completion Method.

The recasting of the Order Completion Method in the setting of uniform convergence spaces, instead of that of ordered sets and their completions, is the first and basic aim of this work. In this regard, appropriate uniform convergence spaces are introduced, and the completions of these spaces are characterized. The existence and uniqueness of generalized solutions of arbitrary continuous systems of nonlinear PDEs is proved within the context of the mentioned uniform convergence spaces. Furthermore, it is shown that these solutions may be assimilated with usual normal lower semi-continuous functions. That is, there is a natural injective and uniformly continuous mapping from the space of generalized solutions into the space of nearly finite normal lower semi-continuous maps. This provides a blanket regularity for the solutions.

The regularity of the generalized solutions delivered through the Order Completion Method is dramatically improved upon in two ways. In the first place, it is shown that such generalized solutions may in fact be assimilated with functions that are smooth, up to the order of differentiability of the nonlinear partial differential operator $T(x, D)$, everywhere except on a closed nowhere dense subset of the domain of definition of the system of equations. This result is based on the fact that any Hausdorff convergence structure admits a complete uniform convergence structure. As such, it seems unlikely that such a result can be obtained in terms of the purely order theoretic methods upon which the Order Completion Method is based.

As mentioned in Section 1.4, the spaces of generalized functions delivered through the Order Completion Method are, to some extent, dependent on the particular nonlinear partial differential operator that defines the equation. For the spaces of generalized functions mentioned above, which are obtained as the Wyler completions of suitable uniform convergence spaces, this is also the case. The second major development we present here, regarding to the regularity of generalized solutions of systems of nonlinear PDEs, addresses this issue. In this regard, and in the original spirit of Sobolev, we construct spaces of generalized functions that do not depend in any way on a particular nonlinear partial differential operator. These spaces are shown to contain generalized solutions, in a suitable sense, of a large class of systems of nonlinear PDEs. This result provides some additional insight into the structure of unique generalized solutions constructed in the Order Completion Method. In particular, such a solution may be interpreted as nothing but the set of solutions in the new Sobolev type spaces of generalized functions.

The solutions constructed in the Sobolev type spaces of generalized functions may be represented through their generalized partial derivatives as usual nearly finite normal lower semi-continuous functions. As such, the singularity set of such a generalized function, that is, the set of points where at least one of the generalized partial derivatives is not continuous, is a set of first Baire category. However, it should be noted that the generalized derivatives cannot be interpreted classically, that is, as usual partial derivatives, at those points where the generalized function is regular.

In this regard, we show that, for a large class of equations, there are generalized solutions which are in fact *classical* solutions everywhere except on some closed nowhere dense set. This result is based on a suitable approximation of functions

$$u : \Omega \rightarrow \mathbb{R}$$

that are $\mathcal{C}^m$-smooth everywhere except on a closed nowhere dense subset of Ω, by functions in $\mathcal{C}^m(\Omega)$, and on a result giving sufficient conditions for the compactness of a set in $\mathcal{C}^m(\Omega)$ with respect to a suitable topology.

The last topic to be treated concerns initial and / or boundary value problems. The results discussed so far apply to systems of nonlinear PDEs without any additional conditions. In this regard, we show that the methods that have been developed here may be applied to initial and / or boundary value problems with only minimal modifications. In particular, we show that a large class of Cauchy problems admit solutions in the Sobolev type spaces of generalized functions. Furthermore, under only very mild assumptions regarding the smoothness of the nonlinear partial differential operator and the initial data, we show that a solution can be constructed which is in fact a classical solution everywhere except on a closed nowhere dense subset of the domain of definition of the system of equations. This result is *a first in the literature*. In particular, it is the first extension of the Cauchy-Kovalevskaia Theorem 2 on its own general and type independent grounds to equations that are not analytic.

5.2 Arrangement of the Material

The results presented in this work are organized as follows. In Chapter 6 we obtain some preliminary results on the Wyler completion of Hausdorff uniform convergence spaces. In particular, we show that if a uniform convergence space X is a subspace of Y, then the completion $X^\sharp$ of X need not be a subspace of Y. However, the inclusion mapping

$$i : X \to Y$$

extends to an injective, uniformly continuous mapping

$$i^\sharp : X^\sharp \to Y^\sharp.$$

More generally, if X and Y are Hausdorff uniform convergence spaces, and

$$\varphi : X \to Y$$

is a uniformly continuous embedding, then the unique uniformly continuous extension

$$\varphi^\sharp : X^\sharp \to Y^\sharp$$

of φ to $X^\sharp$ is injective, but not necessarily an embedding. Products of uniform convergence structure are shown to be compatible the Wyler completion. In particular, the completion of the product $\prod_{i\in I} X_i$ of a family $(X_i)_{i\in I}$ of Hausdorff uniform convergence spaces is the product $\prod_{i\in I} X_i^\sharp$ of the completions $X_i^\sharp$ of the X_i. These results are used to obtain a description of the completion of a uniform convergence space that is equipped with the initial uniform convergence structure with respect to a family of mappings

$$\varphi_i : X \to X_i,$$

where each X_i is a Hausdorff uniform convergence space. In particular, we show that there is an injective, uniformly continuous mapping

$$\Phi : X^\sharp \to \prod_{i\in I} X_i^\sharp$$

so that $\pi_i \circ \Phi = \varphi_i^\sharp$ for each $i \in I$, with π_i the projection.

Chapter 7 concerns certain spaces of normal lower semi-continuous functions. In particular, we introduce the space $\mathcal{NL}(X)$ of all nearly finite normal lower semi-continuous functions, and the space $\mathcal{ML}(X)$ of nearly finite normal lower semi-continuous functions that are continuous and real valued everywhere except on a closed nowhere dense subset of X. This space also appears in connection with rings of continuous functions and their completions [56]. Some properties of the mentioned classes of functions are investigated. We introduce a uniform convergence structure on $\mathcal{ML}(X)$ in such a way that the induced convergence structure is the order convergence structure (4.8). The Wyler completion of $\mathcal{ML}(X)$ with respect to this uniform convergence structure is obtained as the set $\mathcal{NL}(X)$, equipped with a suitable uniform convergence structure.

The uniform convergence spaces introduced in Chapter 7 form the point of departure for the construction of spaces of generalized functions in Chapter 8. In this regard, we construct the so-called pullback space of generalized functions $\mathcal{NL}_{\mathbf{T}}(\Omega)^K$ associated with a given system nonlinear PDEs

$$\mathbf{T}(x, D)\,\mathbf{u}(x) = \mathbf{f}(x)\,. \tag{5.1}$$

in Section 8.1. In Section 8.2 we introduce the Sobolev type spaces of generalized functions $\mathcal{NL}^m(\Omega)$. These spaces are obtained as the completion of the set

$$\mathcal{ML}^m(\Omega) = \left\{ u \in \mathcal{ML}(\Omega) \left| \begin{array}{l} \exists \quad \Gamma \subset \Omega \text{ closed nowhere dense :} \\ \quad u \in \mathcal{C}^m(\Omega) \end{array} \right. \right\}$$

equipped with a suitable uniform convergence structure. The structure of the generalized functions that are the elements of $\mathcal{NL}^m(\Omega)$ are discussed, as well as the connection with the spaces $\mathcal{NL}_{\mathbf{T}}(\Omega)$. In Section 8.3 we discuss the issue of extending a nonlinear partial differential operator to the Sobolev type spaces of generalized functions $\mathcal{NL}^m(\Omega)$. In this regard, we show how such an operator may be defined on $\mathcal{ML}^m(\Omega)^K$, for a suitable $K \in \mathbb{N}$. It is shown that every such operator is uniformly continuous on $\mathcal{ML}^m(\Omega)^K$, and as such it may be uniquely extended to the space $\mathcal{NL}^m(\Omega)^K$. We also discuss the correspondence between generalized solutions in the pullback type spaces of generalized functions, and the Sobolev type spaces of generalized functions. It is shown that generalized solution to (5.1) in $\mathcal{NL}^m(\Omega)^K$ corresponds to the unique generalized solution in the pullback type space $\mathcal{NL}_{\mathbf{T}}(\Omega)$, should such solutions exist.

Chapter 9 addresses the issue of existence of generalized solutions in the spaces of generalized functions constructed in Chapter 8. In Section 9.1 we introduce certain basic approximation results. These include a multidimensional version of (1.110), as well as a suitable refinement of that result. Section 9.2 contains the first and basic existence and uniqueness result for generalized solutions in the pullback spaces of generalized functions $\mathcal{NL}_{\mathbf{T}}(\Omega)$. This section also includes a detailed investigation of the structure of such generalized solutions. In Section 9.3 we investigate the effects of additional assumptions on the smoothness of the nonlinear partial differential operator $\mathbf{T}$ and the righthand term $\mathbf{f}$ on the regularity of generalized solutions in $\mathcal{NL}_{\mathbf{T}}(\Omega)$. Here we obtain what may be viewed as a maximal regularity result for the solution in pullback type spaces of generalized functions. In particular, it is shown that if the nonlinear partial differential operator $\mathbf{T}$, and the righthand term $\mathbf{f}$ are $\mathcal{C}^k$-smooth, for some $k \in \mathbb{N}$, then the generalized solution in $\mathcal{NL}_{\mathbf{T}}(\Omega)$ may be assimilated with functions in $\mathcal{ML}^k(\Omega)^K$. Section 9.4 contains existence results for generalized solutions of a large class of systems of nonlinear PDEs in the Sobolev type spaces of generalized functions. It is also shown that under additional assumptions on the smoothness of the nonlinear partial differential operator $\mathbf{T}$ and righthand term $\mathbf{f}$, namely, that both are $\mathcal{C}^k$-smooth, we may obtain solutions in $\mathcal{NL}^{m+k}(\Omega)$.

In Chapter 10 we discuss further regularity properties of generalized solutions in Sobolev type spaces of generalized functions. Section 10.1 introduces suitable topologies on the space $\mathcal{C}^m(\Omega)$ which admit convenient conditions for a set $\mathcal{A} \subset \mathcal{C}^m(\Omega)$ to be precompact. In particular, we show that any set $\mathcal{F} \subset \mathcal{C}^{m+1}(\Omega)$ that satisfies

$$\begin{array}{ll} \forall & A \subset \Omega \text{ compact :} \\ \exists & M_A > 0 : \\ \forall & |\alpha| \leq m+1 : \\ & u \in A \Rightarrow |D^\alpha u(x)| \leq M_A,\ x \in A \end{array}$$

is precompact in $\mathcal{C}^m(\Omega)$. This generalizes a well known result for the one dimensional case $\Omega \subseteq \mathbb{R}$, see for instance [49]. Using this result, it is shown in Section 10.2 that a large class of systems of nonlinear PDEs admit a generalized solution in $\mathcal{NL}^m(\Omega)^K$ that is in fact a classical solution everywhere except on a closed nowhere dense subset of Ω.

Chapter 11 is dedicated to the study of a large class of initial value problems. In particular, we extend the Cauchy-Kovalevskaia Theorem 2 to systems of equations, and initial data, which are not analytic. In this regard, it is shown that such an initial value problem admits a generalized solution in $\mathcal{NL}^m(\Omega)^K$ which satisfies the initial condition in a suitable generalized sense. This is achieved through a slight modification of the methods introduced in Chapters 7 through 9. Indeed, the initial value problem is solved by essentially the same techniques that apply to the free problem. It is also shown that such a generalized solution to the Cauchy problem may be constructed which is a classical solution everywhere except on a closed nowhere dense subset of the domain of definition. Furthermore, this solution satisfies the initial condition in the usual sense.

Chapter 12 contains some concluding remarks. In particular, we discuss some of the implications of the results obtained here. Directions for future research are also indicated.