Chapter 1

Introduction

The purpose of this book is to explain some recent results in analysis, which in general terms may be described as follows: Major constructions or operations in analysis are often characterized in a natural and unique way by some very elementary properties, relations or equations which they satisfy.

A simple example on the real line would be the exponential function mapping sums to products.

To describe the basic theme of the book, let us consider the classical *Fourier transform* $\mathcal{F}$ on $\mathbb{R}^n$ given by

$$\mathcal{F}(f)(x) = \int_{\mathbb{R}^n} \exp(-2\pi i \langle x, y \rangle) \ f(y) \ dy$$

on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of "rapidly" decreasing smooth functions $f : \mathbb{R}^n \to \mathbb{C}$. As well-known, $\mathcal{F}$ maps bijectively the Schwartz space onto itself and exchanges products with convolutions. The interesting fact is that these properties (almost) characterize the Fourier transform. As shown by Artstein-Avidan, Faifman and Milman [AFM], any bijective transformation $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ satisfying

$$T(f \cdot g) = T(f) * T(g)$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$ is just a slight modification of the Fourier transform: there exists a diffeomorphism $\omega : \mathbb{R}^n \to \mathbb{R}^n$ such that either $T(f) = \mathcal{F}(f \circ \omega)$ for all $f \in \mathcal{S}(\mathbb{R}^n)$ or $T(f) = \overline{\mathcal{F}(f \circ \omega)}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Assuming in addition that T also maps convolutions into products,

$$T(f * g) = T(f) \cdot T(g)$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$, the diffeomorphism ω is given by a linear map $A \in GL(n, \mathbb{R})$ with $|\det(A)| = 1$, i.e., $T(f) = \mathcal{F}(f \circ A)$ for all $f \in \mathcal{S}(\mathbb{R}^n)$ or $T(f) = \overline{\mathcal{F}(f \circ A)}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. Details are found in the papers by the above authors and Alesker [AAM], [AAFM] and [AFM].

H. König, V. Milman, *Operator Relations Characterizing Derivatives*,
https://doi.org/10.1007/978-3-030-00241-1_1

Note that T is not assumed to be linear or continuous. Nevertheless the real linearity and the continuity of T are a consequence of the result.

A priori the Fourier transform is an analytic construction. However, it is essentially uniquely recovered by the basic properties which we mentioned. Starting with the simple formula for a bijective map exchanging products and convolutions, we get an operation which has a rich structure and extremely useful properties.

There are some other properties which also characterize the Fourier transform, e.g., the Poisson summation formula, cf. Faifman [F1], [F2].

The exponential function $e = \exp : \mathbb{R} \to \mathbb{R}$ on the real line is, up to multiples, characterized by its *functional equation*

$$e(x+y) = e(x) \cdot e(y)$$

for all $x, y \in \mathbb{R}^n$, if measurability of e is assumed, see Aczél [A]. In comparison, the Fourier transform $T = \mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is characterized, up to diffeomorphism and complex conjugation, by being bijective and satisfying the *operator functional equation*

$$T(f \cdot g) = T(f) * T(g)$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$. Therefore we recover a classical transform in analysis by an elementary relation, namely the above operator functional equation. Note that the operator T is not assumed to satisfy *any* regularity condition like continuity or measurability as in the case of the above map e.

Following a similar approach, we will study in this book the question to which extent the derivative is characterized by properties like the *Leibniz rule operator equation*

$$T(f \cdot g) = T(f) \cdot g + f \cdot T(g)$$

or the *chain rule operator equation*

$$T(f \circ g) = T(f) \circ g \cdot T(g)$$

on classical function spaces like the spaces C^k of k-times continuously differentiable functions, $T : C^k \to C$, $f, g \in C^k$. We will determine all solutions of either one equation and also of various extensions of them. In most cases, we will a priori assume neither continuity nor linearity or another algebraic property of the operator T. However, a posteriori, a natural type of continuity of T will be a consequence of the result.

Simple additional *initial* conditions like $T(-2\,\mathrm{Id}) = -2$ in the case of the chain rule will guarantee that T is actually the derivative, $Tf = f'$, and, in particular, linear, see Chapter 4.

Returning to the Fourier transform, suppose that $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is bijective and satisfies $T(f \cdot g) = T(f) * T(g)$. By the properties of the Fourier transform, $J := \mathcal{F} \circ T$ then satisfies $J(f \cdot g) = J(f) \cdot J(g)$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$. To prove the result for T, it therefore suffices to determine all bijective multiplicative maps $J : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, $J(f \cdot g) = J(f) \cdot J(g)$, i.e., solve another simple operator functional equation on a classical function space of analysis. In this case there is a diffeomorphism $\omega : \mathbb{R}^n \to \mathbb{R}^n$ such that either $Jf = f \circ \omega$ for all $f \in \mathcal{S}(\mathbb{R}^n)$ or $Jf = \overline{f \circ \omega}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$, cf. [AFM]. Bijective multiplicative maps on relevant functions spaces of analysis were studied before in the papers of Milgram [M] and of Mrčun and Šemrl [Mr], [MS].

Another transformation which is important in analysis and geometry is the Legendre transform $\mathcal{L}$. Let $\mathcal{C}_n$ denote the class of all lower-semi-continuous functions $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ and fix some scalar product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$. The *Legendre transform* of ϕ, also called *Legendre–Fenchel transform* in higher dimensions, is given by

$$\mathcal{L}(\phi)(x) = \sup[\langle x, y\rangle - \phi(y)], \quad \phi \in \mathcal{C}_n,\ x \in \mathbb{R}^n.$$

Then $\mathcal{L}(\phi) \in \mathcal{C}_n$, $\mathcal{L}$ is an involution , i.e., $\mathcal{L}^2(\phi) = \phi$, and $\mathcal{L}$ is order-reversing, i.e., $\phi \leq \psi$ implies $\mathcal{L}(\phi) \geq \mathcal{L}(\psi)$ for all $\phi, \psi \in \mathcal{C}_n$. Being an involution and order-reversing are the most basic properties of a "duality" relation, which is a natural operation having many other interesting and very useful consequences. In fact, they nearly characterize the Legendre transform. By a result of Artstein-Avidan and Milman [AM], for any order-reversing involution $T : \mathcal{C}_n \to \mathcal{C}_n$ there is a symmetric linear map $B \in GL(n, \mathbb{R})$ and there are $v_0 \in \mathbb{R}^n$ and $c_0 \in \mathbb{R}$ such that T has the form $T(\phi) = \mathcal{L}(\phi \circ B + v_0) + \langle\ \cdot, v_0\rangle + c_0$, $\phi \in \mathcal{C}_n$. So up to affine transformations, T is the Legendre transform.

The general problem considered in this book, whether basic constructions or operations in analysis or geometry are essentially characterized by very simple properties like order-reversion or some functional operator equations, was actually motivated by the question what "duality" or "polarity" means in convex geometry and convex analysis. Let $\mathcal{K}_n$ denote the class of closed convex bodies with 0 in its interior. For $K \in \mathcal{K}_n$, the *polar body* $K^\circ \in \mathcal{K}_n$ is given by

$$K^\circ = \{x \in \mathbb{R}^n \mid \text{ for all } y \in K :\ \langle x, y\rangle \leq 1\}.$$

Then the map $K \mapsto K^\circ$ from $\mathcal{K}_n$ to itself is an involution which is order-reversing, $K \subset L$ implying $K^\circ \supset L^\circ$ for all $K, L \in \mathcal{K}_n$. Gruber [Gr] (in a different language), Böröczky, Schneider [BS] and Artstein-Avidan, Milman [AM] (in different setups) showed that conversely any involution $T : \mathcal{K}_n \to \mathcal{K}_n$ which is order-reversing, $K \subset L$ implying $T(K) \supset T(L)$, is actually the polar map, up to linear transformations: There exists $B \in GL(n, \mathbb{R})$ such that T is given by $T(K) = (B(K))^\circ$ for all $K \in \mathcal{K}_n$. The result for the Legendre transform is a corresponding duality result for convex functions instead of convex bodies.

Let us turn back to analysis. The main attention in this book will be given to the study of properties of the derivative, like the Leibniz or the chain rule, and to the question to what extent any of these operator functional equations will nearly characterize the derivative, or what other solutions they admit. We also consider characterizations of the Laplacian and other second-order derivative operations.

It is interesting to compare these classical operations in analysis with functional equations which the exponential function or the logarithm satisfy. The logarithm $\log : \mathbb{R}^+ \to \mathbb{R}$ sends products to sums, $\log(xy) = \log(x) + \log(y)$ for $x, y \in \mathbb{R}^+$. However, on a linear class of functions an analogous non-trivial operation T satisfying $T(f \cdot g) = T(f) + T(g)$ does not exist: considering $g = 0$, one finds that $Tf = 0$ for all f, i.e., $T = 0$. Let us change this operator equation slightly by allowing some "tuning" operators A_1, A_2 which will act on a larger linear space of functions (of "lower" order). For example, consider $T : C^1 \to C$ and $A_1, A_2 : C \to C$ such that

$$T(f \cdot g) = T(f) \cdot A_1(g) + A_2(f) \cdot T(g) \tag{1.1}$$

for all $f, g \in C^1$. Then T still maps in some sense products to sums, with some correction by the tuning operators A_1 and A_2. Clearly, if $A_1 = A_2 = \mathrm{Id}$, we just get the *Leibniz rule equation*, or simply *Leibniz equation*,

$$T(f \cdot g) = T(f) \cdot g + f \cdot T(g), \;\; f, g \in C^1.$$

In addition to the derivative, $T(f) = f'$, also the entropy operation $T(f) = f \ln |f|$ satisfies this equation in C^1 or C, reflecting a logarithmic behavior. The general solution of the Leibniz rule equation turns out to be a linear combination of the derivative and the entropy operation. We prove this in Chapter 3. In this extended interpretation, the derivative operation is an *analogue of the logarithm on linear spaces of functions*. We also determine the solutions of the more general equation (1.1) in Chapter 3.

Another algebraically inspired, interesting aspect of the derivative is illustrated by the chain rule equation

$$T(f \circ g) = T(f) \circ g \cdot T(g) \tag{1.2}$$

for all $f, g \in C^k(\mathbb{R})$, $k \in \mathbb{N}$. In this case T maps the composition $f \circ g = f(g)$ to a "compound" product $T(f) \circ g \cdot T(g)$. Since the information on the left-hand side of the equation involves the composition $f \circ g$ and not individually f and g, on the right-hand side also the composition with g is needed, when f appears, to yield meaningful solutions. A simple product equation $T(f \circ g) = T(f) \cdot T(g)$ only admits the trivial solutions $T = 0$ and $T = 1$. The solutions of the chain rule operator equation are classified in Chapter 4. There are other solutions besides the derivative (and its powers), but the derivative can be characterized by the chain rule and an additional initial condition, e.g., $T(-2\,\mathrm{Id}) = -2$.

Suppose the terms in the chain rule equation (1.2) for T are positive. Then $P := \log T$ satisfies

$$P(f \circ g) = P(f) \circ g + P(g), \quad f, g \in C^1(\mathbb{R}), \tag{1.3}$$

mapping compositions to a "compound sum". Note that equation (1.3) makes sense for all, and not only for positive functions. Again, compositions with g are needed on both sides when terms with f appear.

On linear classes of functions like $C^1(\mathbb{R})$ or $C(\mathbb{R})$, the solutions of equation (1.3) are easily described: there is a continuous function $H \in C(\mathbb{R})$ such that

$$Pf = H \circ f - H.$$

This solution by itself is not very interesting. So let us add on the right-hand side of (1.3) some "tuning" operators, as we did in (1.1). This yields the operator equation

$$T(f \circ g) = T(f) \circ g \cdot A_1(g) + A_2(f) \circ g \cdot T(g) \tag{1.4}$$

with three operators T, A_1, A_2.

One solution of this equation is well-known, namely the second derivative $T(f) = D^2(f) = f''$, with $A_1(f) = (f')^2$ and $A_2(f) = f'$. Natural domains are $C^2(\mathbb{R})$ for T and $C^1(\mathbb{R})$ for A_1, A_2, so A_1, A_2 may be considered of "lower order". In our interpretation, this second-order chain rule equation appears after a logarithmic operation is applied to the first-order chain rule, which then is appropriately "tuned".

We study the solutions of equation (1.4) in Chapter 9 under a mild condition of non-degeneration, and determine all triples of operators (T, A_1, A_2) on suitable $C^k(\mathbb{R})$-spaces which lead to nontrivial solutions.

The operators (T, A_1, A_2) are intertwined by (1.4), and there are fewer types of solutions than one might imagine at first. There are non-trivial solutions for T on $C^k(\mathbb{R})$-spaces for $k \in \{0, 1, 2, 3\}$, with appropriately chosen tuning operators A_1, A_2. On $C^k(\mathbb{R})$-spaces for $k \geq 4$ there are no further solutions, i.e., solutions which might depend on the fourth or higher derivatives. The only solution for $k = 0$ was already described above, $Tf = H \circ f - H$, with $A_1 = A_2 = \mathbb{1}$.

For $k = 1$ there are three different families of solutions, where all operators act on C^1. For $k = 2$, in addition to the solutions mentioned for $k = 0, 1$, there is very little diversity for the operators A_1, A_2. They are again defined on C^1 with $A_1(f) = f' \cdot A_2(f)$ and $A_2(f) = |f'|^p \, \{\operatorname{sgn} f'\}$ for a suitable $p \geq 1$. The term $\{\operatorname{sgn} f'\}$ may appear here or not, yielding two solutions. The main operator T is described by the above value of p and two continuous parameter functions c, H, $c \neq 0$, namely

$$Tf = (cf'' + [H \circ f - H] \cdot f') \cdot |f'|^{p-1} \, \{\operatorname{sgn} f'\}.$$

So for $H = 0$, we essentially get the second derivative. Suitable additional initial conditions determine the form of T. Requiring, e.g., $T(\mathrm{Id}^2) = 2$ and $T(\mathrm{Id}^3) = 6\,\mathrm{Id}$, with $\mathrm{Id}^l(x) = x^l$, $l \in \mathbb{N}$, yields $T(f) = f''$, $A_1(f) = f'^2$ and $A_2(f) = f'$.
The case $k = 3$, in addition to the solutions for $k = 0, 1, 2$, leads to solutions in terms of the Schwarzian derivative S. In this case $A_1(f) = f'^2 A_2(f)$, where A_2

has the same form as for $k = 2$ but with $p \geq 2$. The most interesting solution is $T(f) = f'^2 S(f)$, $A_1(f) = f'^4$ and $A_2(f) = f'^2$, cf. Chapter 9.

One main step of our method is the localization technique, which allows to reduce operator equations to functional equations. We show, e.g., in the case of the chain rule equation (1.2) that any non-degenerate solution $Tf(x)$ is determined by some function F of the variables x, $f(x)$ and the derivatives of f at x up to order k, if $f \in C^k(\mathbb{R})$ and $x \in \mathbb{R}$. No regularity of this function F is known at the outset but has to be proved later. We show that the operator equation for T then turns into a functional equation for F, the solutions of which have to be determined. Usually then we have to prove the continuity or regularity of the coefficient functions appearing in the structure of the solutions of F. Functional equations and regularity results for them are studied in Chapter 2, in preparation for later application in subsequent chapters.

We already mentioned that various of our results are proved under some condition of non-degeneration. There are two different forms and reasons for this type of assumption.

One of them is a very weak form of surjectivity of the operator. This together with the operator equation will often yield in the final result that T is actually surjective. For example, the assumption in Theorem 4.1 for the chain rule equation only requires as non-degeneration condition that T is not the zero operator on the half-bounded functions, allowing a complete description of all solutions.

A very different type of non-degeneration is required when two or three different operators appear in the equation, such as in (1.1). We then need, e.g., that a tuning operator A will not be proportional on some open interval to the operator T, cf. Theorem 3.7, or not be proportional to the identity, cf. Theorem 7.2. By these conditions of non-degeneration we avoid a "resonance" behavior of two different operators, which often has the consequence that they are not localized. In the case of equation (1.1) there is, e.g., the following non-localized solution

$$T(f)(x) = f(x) - f(x+1), \quad A_1(f)(x) = A_2(f)(x) = \frac{1}{2}(f(x) + f(x+1)),$$

where the operators T, A_1 and A_2 act from $C(\mathbb{R})$ to itself. For functions with small support around x, T here acts as identity and A_1 and A_2 are homothetic to the identity. These effects typically appear with Leibniz rule type equations which are studied in Chapters 3 and 7. The exact form of the non-degeneration condition differs from one chapter to the other, but stays the same in each chapter.

In some cases we may avoid the assumption of non-degeneration and prove theorems about the general structure of the solutions of equations like (1.1) without localization. These results are found in Chapter 8.

Interestingly enough, the equations we consider in this book show some unexpected stability or even rigidity. Perturbing the Leibniz rule equation by a "small" additive term yields solutions which are perturbations of the original equation. In the case of the chain rule, we even have rigidity: the perturbed solutions have the same solutions as the original equation. The chain rule equation allows no reasonable additive perturbation. This is shown in Theorems 5.6 and 5.8.

In the case of the chain rule, the rigidity even allows us to study the solutions of the *inequality*

$$T(f \circ g) \leq T(f) \circ g \cdot T(g), \quad f, g \in C^1(\mathbb{R}).$$

To completely describe the solutions of this operator inequality, under some non-degeneration condition and a weak continuity assumption, we have to prove localization and classify certain submultiplicative functions on the real line, which by itself is a curious result, cf. Theorems 6.1 and 6.2.

Let us consider not necessarily small additive "perturbations" of the Leibniz rule. Suppose, e.g., that we add to the Leibniz equation a product of two copies of a "lower-order" operator A,

$$T(f \cdot g) = Tf \cdot g + f \cdot Tg + Af \cdot Ag,$$

$f, g \in C^k(\mathbb{R})$. This equation is not only motivated by a perturbation of the (first order) Leibniz rule, but, in fact, reflects the behavior of the second derivative $T = D^2$. Indeed, choosing $A = \sqrt{2}\, D$, the equation is satisfied for these operators (T, A). The natural domain for T is $C^2(\mathbb{R})$, for A it is $C^1(\mathbb{R})$. Thus A is of "lower order" than T.

This point of view leads to higher-order Leibniz rule type equations determining derivatives of any order, cf. Section 3 of Chapter 5. Moreover, it may be considered for functions on $\mathbb{R}^n$, too. The equations then yield characterizations of the Laplacian under natural assumptions, e.g., orthogonal invariance and annihilation of affine functions. We investigate this in Chapter 7.

In most of the results on operator equations for one operator T in this book we do not make any continuity or regularity assumption on the operator T. A posteriori, the theorems imply that the operator T is actually continuous in a natural way. In the proofs we use the fact that the image of the operators is contained in spaces of continuous functions. We feel, however, that the main reason for the automatic continuity of the solution operators is a consequence of the *non-linearity* of the equations, like the chain rule equation. Using the axiom of choice and Hamel bases, it is of course easy to construct non-continuous and even non-measurable solutions of *linear* equations on infinite-dimensional spaces. However, this is not the case for non-linear equations.

Much of the material of the book is based on papers of the authors and their coauthors. However, various theorems shown in this book extend published results or relax the assumptions made there. The proofs of most results are provided in detail. The book is addressed to a general mathematical audience.

We are very grateful to Mrs. M. Hercberg for her invaluable help in typing and editing the text. The book could not have appeared without her efforts. We also thank several colleagues for helpful discussions on various aspects of the book, in particular S. Alesker, S. Artstein-Avidan, J. Bernstein, P. Domański, D. Faifman, R. Farnsteiner and L. Polterovich.