Pseudo-differential Operators and Symmetries

M. Ruzhansky and V. Turunen

July 22, 2009

Contents

Part II Commutative Symmetries

Preface

This monograph is devoted to the development of the theory of pseudo-differential operators on spaces with symmetries. Such spaces are the Euclidean space $\mathbb{R}^n$, the torus $\mathbb{T}^n$, compact Lie groups and compact homogeneous spaces.

The book consists of several parts. One of our aims has been not only to present new results on pseudo-differential operators but also to show parallels between different approaches to pseudo-differential operators on different spaces. Moreover, we tried to present the material in a self-contained way to make it accessible for readers approaching the material for the first time.

However, different spaces on which we develop the theory of pseudo-differential operators require different backgrounds. Thus, while operators on the Euclidean space in Chapter 2 rely on the well-known Euclidean Fourier analysis, pseudo-differential operators on the torus and more general Lie groups in Chapters 4 and 10 require certain backgrounds in discrete analysis and in the representation theory of compact Lie groups, which we therefore present in Chapter 3 and in Part III, respectively. Moreover, anyone who wishes to work with pseudo-differential operators on Lie groups will certainly benefit from a good grasp of certain aspects of representation theory. That is why we present the main elements of this theory in Part III, thus eliminating the necessity for the reader to consult other sources for most of the time. Similarly, the backgrounds for the theory of pseudo-differential operators on $\mathbb{S}^3$ and $\mathrm{SU}(2)$ developed in Chapter 12 can be found in Chapter 11 presented in a self-contained way suitable for immediate use.

However, it was still not a simple matter to make a self-contained presentation of these theories without referring to basics of the more general analysis. Thus, in hoping that this monograph may serve as a guide to different aspects of pseudo-differential operators, we decided to include the basics of analysis that are certainly useful for anyone working with pseudo-differential operators.

Overall, we tried to supplement all the material with exercises for learning the ideas and practicing the techniques. They range from elementary problems to more challenging ones. In fact, on many occasions where other authors could say "it is easy to see" or "one can check", we prefer to present it as an exercise. At the same time, more challenging exercises also serve as an excellent way to present more aspects of the discussed material.

We would like to thank Professor G. Vainikko, who introduced V. Turunen to pseudo-differential equations on circles [136], leading naturally to the non-commutative setting of the doctoral thesis. The thesis work was crucially influenced by a visit to M.E. Taylor in spring 2000. We are grateful to Professor M.W. Wong for suggesting that we write this monograph, to our students for giving us useful feedback on the background material of the book, and to Dr. J. Wirth for reading the manuscript and for his useful feedback and numerous comments, which led to clarifications of the presentation, especially of the material from Section 10.3. Most of the work was carried out at the pleasant atmospheres provided by Helsinki University of Technology and Imperial College London. Moreover, over the years, we have outlined substantial parts of the monograph elsewhere: particularly, we appreciate the hospitality of University of North Carolina at Chapel Hill, University of Torino and Osaka University. The work of M. Ruzhansky was supported in part by EPSRC grants EP/E062873/01 and EP/G007233/1. The travels of V. Turunen were financed by the Magnus Ehrnrooth Foundation, by the Vilho, Yrjö and Kalle Väisälä Foundation of the Finnish Academy of Science and Letters, and by the Finnish Cultural Foundation. Finally, our loving thanks go to our families for all the encouragement and understanding that we received while working on this monograph.

March 2009

Michael Ruzhansky, London
Ville Turunen, Helsinki

Introduction

Historical notes

Pseudo-differential operators (ΨDO) can be considered as natural extensions of linear partial differential operators, with which they share many essential properties. The study of pseudo-differential operators grew out of research in the 1960s on singular integral operators; being a relatively young subject, the theory is only now reaching a stable form.

Pseudo-differential operators are generalisations of linear partial differential operators, with roots entwined deep down in solving differential equations.

Among the most influential predecessors of the theory of pseudo-differential operators one must mention the works of Solomon Grigorievich Mikhlin, Alberto Calderón and Antoni Szczepan Zygmund. Around 1957, anticipating novel methods, Alberto Calderón proved the local uniqueness theorem of the Cauchy problem of a partial differential equation. This proof involved the idea of studying the algebraic theory of characteristic polynomials of differential equations.

Another landmark was set in ca. 1963, when Michael Atiyah and Isadore Singer presented their celebrated index theorem. Applying operators, which nowadays are recognised as pseudo-differential operators, it was shown that the geometric and analytical indices of Erik Ivar Fredholm's "Fredholm operator" on a compact manifold are equal. In particular, these successes by Calderón and Atiyah–Singer motivated developing a comprehensive theory for these newly found tools. The Atiyah–Singer index theorem is also tied to the advent of K-theory, a significant field of study in itself.

The evolution of the pseudo-differential theory was then rapid. In 1963, Peter Lax proposed some singular integral representations using Jean Baptiste Joseph Fourier's "Fourier series". A little later, Joseph Kohn and Louis Nirenberg presented a more useful approach with the aid of Fourier integral operators and named their representations pseudo-differential operators. Showing that these operators form an algebra, they derived a broad theory, and their results were applied by Peter Lax and Kurt Otto Friedrichs in boundary problems of linear partial differential equations. Other related studies were conducted by Agranovich, Bokobza, Kumano-go, Schwartz, Seeley, Unterberger, and foremost, by Lars Hörmander, who coined the modern pseudo-differential theory in 1965, leading into a vast range of methods and results. The efforts of Kohn, Nirenberg, and Hörmander gave birth to symbol analysis, which is the basis of the theory of pseudo-differential operators.

It is interesting how the ideas of symbol analysis have matured over about 200 years. Already Joseph Lagrange and Augustin Cauchy studied the assignment of a characteristic polynomial to the corresponding differential operator. In the 1880s, Oliver Heaviside developed an operational calculus for the solution of ordinary differential equations met in the theory of electrical circuits. A more sophisticated problem of this kind, related to quantum mechanics, was solved by Hermann Weyl in 1927, and eventually the concept of the symbol of an operator was introduced by Solomon Grigorievich Mikhlin in 1936. After all, there is nothing new under the sun.

Since the mid-1960s, pseudo-differential operators have been widely applied in research on partial differential equations: along with new theorems, they have provided a better understanding of parts of classical analysis including, for instance, Sergei Lvovich Sobolev's "Sobolev spaces", potentials, George Green's "Green functions", fundamental solutions, and the index theory of elliptic operators. Furthermore, they appear naturally when reducing elliptic boundary value problems to the boundary. Briefly, modern mathematical analysis has gained valuable clarity with the unifying aid of pseudo-differential operators. Fourier integral operators are more general than pseudo-differential operators, having the same status in the study of hyperbolic equations as pseudo-differential operators have with respect to elliptic equations.

A natural approach to treat pseudo-differential operators on n-dimensional C^∞-manifolds is to use the theory of $\mathbb{R}^n$ locally: this can be done, since the classes of pseudo-differential operators are invariant under smooth changes of coordinates. However, on periodic spaces (tori) $\mathbb{T}^n$, this could be a clumsy way of thinking, as the local theory is plagued with rather technical convergence and local coordinate questions. The compact group structure of the torus is important from the harmonic analysis point of view.

In 1979 (and 1985) Mikhail Semenovich Agranovich (see [3]) presented an appealing formulation of pseudo-differential operators on the unit circle $\mathbb{S}^1$ using Fourier series. Hence, the independent study of periodic pseudo-differential operators was initiated. The equivalence of local and global definitions of periodic pseudo-differential operators was completely proven by William McLean in 1989. By then, the global definition was widely adopted and used by Agranovich, Amosov, D.N. Arnold, Elschner, McLean, Saranen, Schmidt, Sloan, and Wendland among others. Its effectiveness has been recognised particularly in the numerical analysis of boundary integral equations.

The literature on pseudo-differential operators is extensive. At the time of writing of this paragraph (28 January 2009), a search on MathSciNet showed 1107 entries with words "pseudodifferential operator" in the title (among which 33 are books), 436 entries with words "pseudo-differential operator" in the title (among which 37 are books), 3971 entries with words "pseudodifferential operator" anywhere (among which 417 are books), and 1509 entries with words "pseudo-differential operator" anywhere (among which 151 are books). Most of these works are devoted to the analysis on $\mathbb{R}^n$ and thus we have no means to give a comprehensive overview there. Thus, the emphasis of this monograph is on pseudo-differential operators on the torus, on Lie groups, and on spaces with symmetries, in which cases the literature is much more limited.

Periodic pseudo-differential operators

It turns out that the pseudo-differential and periodic pseudo-differential theories are analogous, the periodic case actually being more discernible.

Despite the intense research on periodic integral equations, the theory of periodic pseudo-differential operators has been difficult to find in the literature. On the other hand, the wealth of publications on general pseudo-differential operators is cumbersome for the periodic case, and it is too easy to get lost in the midst of irrelevant technical details.

In the sequel the elementary properties of periodic pseudo-differential operators are studied. The prerequisites for understanding the theory are more modest than one might expect. Of course, a basic knowledge of functional analysis is necessary, but the simple central tools are Gottfried Wilhelm von Leibniz' "Leibniz formula", Brook Taylor's "Taylor expansion", and Jean Baptiste Joseph Fourier's "Fourier transform". In the periodic case, these familiar concepts of the classical calculus are to be expressed in discrete forms using differences and summation instead of derivatives and integration.

Our working spaces will be the Sobolev spaces $H^s(\mathbb{T}^n)$ on the compact torus group $\mathbb{T}^n$. These spaces ideally reflect smoothness properties, which are of fundamental significance for pseudo-differential operators, as the traditional operator theoretic methods fail to be satisfactory – pseudo-differential operators and periodic pseudo-differential operators do not form any reasonable normed algebra.

The structure of the treatment of periodic pseudo-differential operators is the following: first, introduction of necessary functional analytic prerequisites, then development of useful tools for analysis of series and periodic functions, and after that the presentation of the theory of periodic pseudo-differential operators. The focus of the study is on symbolic analysis.

The techniques of the extension of symbols and the periodisation of operators allow one effectively to relate the Euclidean and the periodic theories, and to use one to derive results in the other. However, we tried to reduce a reliance on such ideas, keeping in mind the development of the subject on Lie groups where such a relation is not readily available. From this point of view, analysis on the torus can be viewed rather as a special case of analysis on a Lie group than the periodic Euclidean case.

The main justification of this work on the torus, from the authors' point of view, is the unification and development of the global theory of periodic pseudo-differential operators. It becomes evident how elegant this theory is, especially when compared to the theory on $\mathbb{R}^n$; and as such, periodic pseudo-differential operators may actually serve as a nice first introduction to the general theory of pseudo-differential operators. For those who have already acquainted themselves with pseudo-differential operators this work may still offer another aspect of the analysis. Thus, there is a hope that these tools will find various uses.

Although we decided not to discuss Fourier integral operators on $\mathbb{R}^n$, we devote some efforts to analysing operators that we call Fourier series operators. These are analogues of Fourier integral operators on the torus and we study them in terms of toroidal quantization. The main new difficulty here is that while pseudo-differential operators do not move the wave front sets of distributions, this is no longer the case for Fourier series operators. Thus, we are quickly forced to make extensions of functions from an integer lattice to Euclidean space on the frequency side. The analysis presented here shows certain limitation of the use of Fourier series operators; however, we succeed in establishing elements of calculus for them and discuss an application to hyperbolic partial differential equations.

Pseudo-differential operators on Lie groups

Non-commutative Lie groups and homogeneous spaces play important roles in different areas of mathematics. Some fundamental examples include spheres $\mathbb{S}^n$, which are homogeneous spaces under the action of the orthogonal groups. The important special case is the three-dimensional sphere $\mathbb{S}^3$ which happens to be also a group. However, while the general theory of pseudo-differential operators is available on such spaces, it presents certain limitations. First, working in local coordinates often makes it very complicated to keep track of the global geometric features. For example, a fundamental property that spheres are fixed by rotations becomes almost untraceable when looking at it in local coordinates. Another limitation is that while the local approach yields an invariant notion of the principal symbol, the full symbol is not readily available. This presents profound complications in applying the theory of pseudo-differential operators to problems on manifolds that depend on knowledge of the full symbol of an operator.

In general, it is a natural idea to build pseudo-differential operators out of smooth families of convolution operators on Lie groups. There have been many works aiming at the understanding of pseudo-differential operators on Lie groups from this point of view, e.g., the works on left-invariant operators [120, 78, 40], convolution calculus on nilpotent Lie groups [77], L^2-boundedness of convolution operators related to Howe's conjecture [57, 41], and many others.

However, in this work, we strive to develop the convolution approach into a symbolic quantization, which always provides a much more convenient framework for the analysis of operators. For this, our analysis of operators and their symbols is based on the representation theory of Lie groups. This leads to a description of the full symbol of a pseudo-differential operator on a Lie group as a sequence of matrices of growing sizes equal to the dimensions of the corresponding representations of the group. We also characterise, in terms of the introduced quantizations, standard Hörmander classes Ψ^m on Lie groups. One of the advantages of the presented approach is that we obtain a notion of full (global) symbols which matches the underlying Fourier analysis on the group in a perfect way. For a group G, such a symbol can be interpreted as a mapping defined on the space $G \times \widehat{G}$, where $\widehat{G}$ is the unitary dual of a compact Lie group G. In a nutshell, this analysis can be regarded as a non-commutative analogue of the Kohn–Nirenberg quantization of pseudo-differential operators that was proposed by Joseph Kohn and Louis Nirenberg in [68] in the Euclidean setting. As such, the present research is perhaps most closely related to the work of Michael Taylor [127], who, however, in his analysis used an exponential mapping to rely on pseudo-differential operators on a Lie algebra which can be viewed as a Euclidean space with the corresponding standard theory of pseudo-differential operators. However, the approach developed in this work is different from that of [127, 128] in the sense that we rely on the group structure directly and thus are not restricted to neighbourhoods of the neutral element, thus being able to approach global symbol classes directly. Some aspects of the analysis presented in this part appeared in [98].

As an important example, the approach developed here gives us quite detailed information on the global quantization of operators on the three-dimensional sphere $\mathbb{S}^3$. More generally, we note that if we have a closed simply-connected three-dimensional manifold M, then by the recently resolved Poincaré conjecture there exists a global diffeomorphism $M \simeq \mathbb{S}^3 \simeq \mathrm{SU}(2)$ that turns M into a Lie group with a group structure induced by $\mathbb{S}^3$ (or by $\mathrm{SU}(2)$). Thus, we can use the approach developed for $\mathrm{SU}(2)$ to immediately obtain the corresponding global quantization of operators on M with respect to this induced group product. In fact, all the formulae remain completely the same since the unitary dual of $\mathrm{SU}(2)$ (or $\mathbb{S}^3$ in the quaternionic $\mathbb{R}^4$) is mapped by this diffeomorphism as well. An interesting feature of the pseudo-differential operators from Hörmander's classes Ψ^m on these spaces is that they have matrix-valued full symbols with a remarkable rapid off-diagonal decay property.

We also introduce a general machinery with which we obtain global quantization on homogeneous spaces using the one on the Lie group that acts on the space. Although we do not yet have general analogues of the diffeomorphic Poincaré conjecture in higher dimensions, this already covers cases when M is a convex surface or a surface with positive curvature tensor, as well as more general manifolds in terms of their Pontryagin class, etc.

Conventions

Each part or a chapter of the book is preceded by a short introduction explaining the layout and conventions. However, let us mention now several conventions that hold throughout the book.

Constants will be usually denoted by C (sometimes with subscripts), and their values may differ on different occasions, even when appearing in subsequent estimates. Throughout the book, the notation for the Laplace operator is $\mathcal{L}$ in order not to confuse it with difference operators which are denoted by $\triangle$.

In Chapters 3 and 4 we encounter a notational difficulty that both frequencies and multi-indices are integers with different conventions for norms than are normally used in the literature. To address this issue, there we let $|\alpha| = |\alpha|_{\ell^1}$ be the ℓ^1-norm (of the multi-index α) and $\|\xi\| = \|\xi\|_{\ell^2}$ be the Euclidean ℓ^2 norm (of the frequency $\xi \in \mathbb{Z}^n$). However, in other chapters we write a more traditional $|\xi|$ for the length of the vector ξ in $\mathbb{R}^n$, and reserve the notation $\|\cdot\|_X$ for a norm in a normed space X. However, there should be no confusion with this notation since we usually make it clear which norm we use. In Part IV, $\xi = \xi(x)$ stands for a representation, so that we can still use the usual notation $\sigma(x, \xi)$ for symbols.

Part I

Foundations of Analysis

Part I of the monograph contains preliminary material that could be useful for anyone working in the theory of pseudo-differential operators.

The material of the book is on the intersection of classical analysis with the representation theory of Lie groups. Aiming at making the presentation self-sufficient we include preliminary material that may be used as a reference for concepts developed later. In any case, the material presented in this part may be used either as a reference or as an independent textbook on the foundations of analysis.

Throughout the book, we assume that the reader has survived undergraduate calculus courses, so that concepts like *partial derivatives* and the *Riemann integral* are familiar. Otherwise, the prerequisites for understanding the material in this book are quite modest. We shall start with a naive version of a set theory, metric spaces, topology, functional analysis, measure theory and integration in Lebesgue's sense.

Chapter A

Sets, Topology and Metrics

First, we present the basic notations and properties of sets, used elsewhere in the book. The set theory involved is "naive", sufficient for our purposes; for a thorough treatment, see, e.g., [46]. The sets of integer, rational, real or complex numbers will be taken for granted, we shall not construct them.

Let us first list some abbreviations that we are going to use:

- "P *and* Q" means that both properties P and Q are true.
- "P *or* Q" means that at least one of the properties P and Q is true.
- "$P \Rightarrow Q$" reads "*If* P *then* Q", meaning that "P is false or Q is true". Equivalently "$Q \Leftarrow P$", i.e., "Q *only if* P".
- "$P \iff Q$" is "$P \Rightarrow Q$ *and* $P \Leftarrow Q$", reading "P *if and only if* Q".
- "$\exists x$" reads "*There exists* x".
- "$\exists! x$" reads "*There exists a unique* x".
- "$\forall x$" reads "*For every* x".
- "$P := Q$" or "$Q =: P$" reads "P *is defined to be* Q".

A.1 Sets, collections, families

Naively, a *set* (or a *collection* or a *family*) A consists of *points* (or *elements* or *members*) x.

Example. Sets of points, like a *collection of coins*, a *family of two parents and three children*, a *flock of sheep*, a *pack of wolves*, or a *crowd of protesters*.

Example. Points in a set, like the *members of a parliament*, the *flowers in a bundle*, or the *stars in a constellation*.

We denote $x \in A$ if the element x *belongs to the set* A, and $x \notin A$ if x does not belong to A. A set A is a *subset* of a set B, denoted by $A \subset B$ or $B \supset A$, if

$$\forall x: \ x \in A \Rightarrow x \in B.$$

Sets A, B are *equal*, denoted by $A = B$, if $A \subset B$ and $B \subset A$, i.e.,

$$\forall x : \ x \in A \iff x \in B.$$

If $A \subset B$ and $A \neq B$ then A is called a *proper subset* of B.

Remark A.1.1 **(Notation for numbers)**. The sets of *integer, rational, real and complex numbers* are respectively $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$; let $\mathbb{N} = \mathbb{Z}^+$ and $\mathbb{R}^+$ stand for the corresponding subsets of (strictly) positive numbers. Then

$$\mathbb{Z}^+ \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

We also write $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

There are various ways for expressing sets. Sometimes all the elements can be listed:

- The *empty set* $\emptyset = \{\}$ is the unique set without elements: $\forall x : x \notin \emptyset$.
- Set $\{x\}$ consists of a single element $x \in \{x\}$.
- Set $\{x, y\} = \{y, x\}$ consists of elements x and y. And so on. Yet $\{x\} = \{x, x\} = \{x, x, x\}$ etc.

A set consisting of those elements for which property P holds can be denoted by

$$\{x : \ P(x)\} = \{x \mid P(x)\}.$$

A set consisting of finitely many elements $x_1, \ldots, x_n$ could be denoted by

$$\begin{aligned} \{x_1, \ldots, x_n\} &= \{x_k : \ k \in \{1, \ldots, n\}\} \\ &= \left\{x_k \mid k \in \mathbb{Z}^+ : \ k \leq n\right\} \\ &= \{x_k\}_{k=1}^n, \end{aligned}$$

and the infinite set of positive integers by

$$\mathbb{Z}^+ = \{1, 2, 3, 4, 5, \cdots\}.$$

The *power set* $\mathcal{P}(X)$ consists of all the subsets of X,

$$\mathcal{P}(X) = \{A : \ A \subset X\}$$

Example. For the set $X = \{1\}$, we have

$$\begin{aligned} \mathcal{P}(X) &= \{\emptyset, \{1\}\}, \\ \mathcal{P}(\mathcal{P}(X)) &= \{\emptyset, \{\emptyset\}, \{1\}, \{\emptyset, \{1\}\}\}, \end{aligned}$$

and we leave it as an exercise to find $\mathcal{P}(\mathcal{P}(\mathcal{P}(X)))$, which contains $2^4 = 16$ elements in this case.

Example. Always at least $\emptyset, X \in \mathcal{P}(X)$. If $x \in X$, then $\{x\} \in \mathcal{P}(X)$ and $\{\{x\}\} \in \mathcal{P}(\mathcal{P}(X))$,

$$x \neq \{x\} \neq \{\{x\}\} \neq \cdots ,$$
$$x \in \{x\} \in \{\{x\}\} \in \cdots .$$

However, we shall allow neither $x \in x$ nor $x \notin x$; consider *Russell's paradox*: given $x = \{a : a \notin a\}$, is $x \in x$?

For $A, B \subset X$, let us define the *union* $A \cup B$, the *intersection* $A \cap B$ and the *difference* $A \setminus B$ by

$$\begin{aligned} A \cup B &:= \{x : \ x \in A \text{ or } x \in B\}, \\ A \cap B &:= \{x : \ x \in A \text{ and } x \in B\}, \\ A \setminus B &:= \{x : \ x \in A \text{ and } x \notin B\}. \end{aligned}$$

The *complement* A^c of A in X is defined by $A^c := X \setminus A$.

Example. If $A = \{1, 2\}$ and $B = \{2, 3\}$ then $A \cup B = \{1, 2, 3\}$, $A \cap B = \{2\}$ and $A \setminus B = \{1\}$.

Example. $\mathbb{R} \setminus \mathbb{Q}$ is the set of irrational numbers.

Exercise A.1.2. Show that

$$\begin{aligned} (A \cup B) \cup C &= A \cup (B \cup C), \\ (A \cap B) \cap C &= A \cap (B \cap C), \\ (A \cup B) \cap C &= (A \cap C) \cup (B \cap C), \\ (A \cap B) \cup C &= (A \cup C) \cap (B \cup C). \end{aligned}$$

Notice that in the latter two cases above, the order of the parentheses is essential. On the other hand, the associativity in the first two equalities allows us to abbreviate $A \cup B \cup C := (A \cup B) \cup C$ and $A \cap B \cap C := (A \cap B) \cap C$ and so on.

Definition A.1.3 (Index sets). Let I be any set and assume that for every $i \in I$ we are given a set A_i. Then I is an *index set* for the collection of sets A_i.

Definition A.1.4 (Unions and intersections of families). For a family $\mathcal{A} \subset \mathcal{P}(X)$, the *union* $\bigcup \mathcal{A}$ and the *intersection* $\bigcap \mathcal{A}$ are defined by

$$\begin{aligned} \bigcup \mathcal{A} = \bigcup_{B \in \mathcal{A}} B &:= \{x \mid \exists B \in \mathcal{A} : \ x \in B\}, \\ \bigcap \mathcal{A} = \bigcap_{B \in \mathcal{A}} B &:= \{x \mid \forall B \in \mathcal{A} : \ x \in B\}. \end{aligned}$$

Example. If $\mathcal{A} = \{B, C\}$ then $\bigcup \mathcal{A} = B \cup C$ and $\bigcap \mathcal{A} = B \cap C$.

Notice that if $\mathcal{A} \subset \mathcal{B} \subset \mathcal{P}(X)$ then

$$\emptyset \subset \bigcup \mathcal{A} \subset \bigcup \mathcal{B} \subset X \quad \text{and} \quad \emptyset \subset \bigcap \mathcal{B} \subset \bigcap \mathcal{A} \subset X.$$

Especially, for $\emptyset \subset \mathcal{P}(X)$ we have

$$\bigcup \emptyset = \emptyset \quad \text{and} \quad \bigcap \emptyset = X. \tag{A.1}$$

Notice that $A \cup B = \bigcup\{A, B\}$ and $A \cap B = \bigcap\{A, B\}$. For unions (and similarly for intersections), the following notations are also commonplace:

$$\begin{aligned}
\bigcup_{j \in K} A_j &:= \bigcup\{A_j \mid j \in K\}, \\
\bigcup_{k=1}^{n} A_k &:= \bigcup\{A_k \mid k \in \mathbb{Z}^+ : \ 1 \leq k \leq n\}, \\
\bigcup_{k=1}^{\infty} A_k &:= \bigcup\{A_k \mid k \in \mathbb{Z}^+\}.
\end{aligned}$$

Example. $\displaystyle\bigcap_{k=1}^{3} A_k = A_1 \cap A_2 \cap A_3.$

Exercise A.1.5 (de Morgan's rules). Prove *de Morgan's rules*:

$$\begin{aligned}
X \setminus \bigcup_{j \in K} A_j &= \bigcap_{j \in K} (X \setminus A_j), \\
X \setminus \bigcap_{j \in K} A_j &= \bigcup_{j \in K} (X \setminus A_j).
\end{aligned}$$

A.2 Relations, functions, equivalences and orders

The *Cartesian product* of sets A and B is

$$A \times B = \{(x, y) : \ x \in A, \ y \in B\},$$

where the elements $(x, y) := \{x, \{x, y\}\}$ are ordered pairs: if $x \neq y$ then $(x, y) \neq (y, x)$, whereas $\{x, y\} = \{y, x\}$. A *relation from A to B* is a subset $R \subset A \times B$. We write xRy if $(x, y) \in R$, saying "*x is in relation R to y*"; analogously, $x \not{R}\, y$ means $(x, y) \notin R$ ("*x is not in relation R to y*").

Functions. A relation $f \subset X \times Y$ is called a *function* (or a *mapping*) from X to Y, denoted by

$$f : X \to Y \quad \text{or} \quad X \xrightarrow{f} Y,$$

if for each $x \in X$ there exists a unique $y \in Y$ such that $(x, y) \in f$:

$$\forall x \in X \; \exists! y \in Y : \quad (x, y) \in f;$$

in this case, we write

$$y := f(x) \quad \text{or} \quad x \mapsto f(x) = y.$$

Intuitively, a function $f : X \to Y$ is a rule taking $x \in X$ to $f(x) \in Y$. Functions $X \xrightarrow{f} Y$ and $Y \xrightarrow{g} Z$ yield a *composition* $X \xrightarrow{g \circ f} Z$ by $g \circ f(x) := g(f(x))$. The *restriction* of $f : X \to Y$ to $A \subset X$ is $f|_A : A \to Y$ defined by $f|_A(x) := f(x)$.

Example. The *characteristic function* of a set $E \in \mathcal{P}(X)$ is $\chi_E : X \to \mathbb{R}$ defined by

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Definition A.2.1 (Injections, surjections, bijections). A function $f : X \to Y$ is

- an *injection* if $f(x_1) = f(x_2)$ implies $x_1 = x_2$,
- a *surjection* if for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$,
- and a *bijection* if it is both injective and surjective, and in this case we may define the *inverse function* $f^{-1} : Y \to X$ such that $f(x) = y$ if and only if $x = f^{-1}(y)$.

Definition A.2.2. (**Image and preimage**) A function $f : X \to Y$ begets functions

$$\begin{aligned} f^+ &: \mathcal{P}(X) \to \mathcal{P}(Y), \quad f^+(A) = f(A) := \{f(x) \in Y : x \in A\}, \\ f^- &: \mathcal{P}(Y) \to \mathcal{P}(X), \quad f^-(B) = f^{-1}(B) := \{x \in Y : f(x) \in B\}. \end{aligned}$$

Sets $f(A)$ and $f^{-1}(B)$ are called the *image* of $A \subset X$ and the *preimage* of $B \subset Y$, respectively.

Exercise A.2.3. Let $f : X \to Y$, $A \subset X$ and $B \subset Y$. Show that

$$A \subset f^{-1}(f(A)) \quad \text{and} \quad f(f^{-1}(B)) \subset B.$$

Give examples showing that these subsets can be proper.

Exercise A.2.4. Let $f : X \to Y$, $A_0 \subset X$, $B_0 \subset Y$, $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$. Show that

$$\begin{cases} f(\bigcup \mathcal{A}) & = & \bigcup_{A \in \mathcal{A}} f(A), \\ f(\bigcap \mathcal{A}) & \subset & \bigcap_{A \in \mathcal{A}} f(A), \\ f(X \setminus A_0) & \supset & Y \setminus f(A_0), \end{cases}$$

where the subsets can be proper, while

$$\begin{cases} f^{-1}(\bigcup \mathcal{B}) & = & \bigcup_{B \in \mathcal{B}} f^{-1}(B), \\ f^{-1}(\bigcap \mathcal{B}) & = & \bigcap_{B \in \mathcal{B}} f^{-1}(B), \\ f^{-1}(Y \setminus B_0) & = & X \setminus f^{-1}(B_0). \end{cases}$$

These set-operation-friendly properties of $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ will be encountered later in topology and measure theory.

Definition A.2.5 (Induced and co-induced families). Let $f : X \to Y$, $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$. Then f is said to *induce* the family $f^{-1}(\mathcal{B}) \subset \mathcal{P}(X)$ and to *co-induce* the family $\mathcal{D} \subset \mathcal{P}(Y)$, where

$$\begin{aligned} f^{-1}(\mathcal{B}) &:= \left\{ f^{-1}(B) \mid B \in \mathcal{B} \right\}, \\ \mathcal{D} &:= \left\{ B \subset Y \mid f^{-1}(B) \in \mathcal{A} \right\}. \end{aligned}$$

Equivalences

Definition A.2.6 (Equivalence relation). A subset $\sim$ of $X \times X$ is an *equivalence relation* on X if it is

1. *reflexive*: $x \sim x$ (for all $x \in X$);
2. *symmetric*: if $x \sim y$ then $y \sim x$ (for all $x, y \in X$);
3. *transitive*: if $x \sim y$ and $y \sim z$ then $x \sim z$ (for all $x, y, z \in X$).

The *equivalence class* of $x \in X$ is

$$[x] := \{y \in X \mid x \sim y\},$$

and the equivalence classes form the *quotient space*

$$X/\sim \quad := \quad \{[x] \mid x \in X\}.$$

Notice that $x \in [x] \subset X$, that $[x] \cap [y] = \emptyset$ if $[x] \neq [y]$, and that $X = \bigcup_{x \in X} [x]$.

Example. Clearly, the identity relation $=$ is an equivalence relation on X, and $f(x) := \{x\}$ defines a natural bijection $f : X \to X/=$.

Example. Let X and Y denote the sets of all women and men, respectively. For simplicity, we may assume the disjointness $X \cap Y = \emptyset$. Let $Isolde, Juliet \in X$ and $Romeo, Tristan \in Y$. For $a, b \in X \cup Y$, let $x \sim y$ if and only if a and b are of the same gender. Then

$$\begin{aligned} Y = [Tristan] = [Romeo] &\neq [Juliet] = [Isolde] = X, \\ X \cup Y &= [Romeo] \cup [Juliet], \\ (X \cup Y)/\sim &= \{[Romeo], [Juliet]\}. \end{aligned}$$

Exercise A.2.7. Let us define a relation $\sim$ in the Euclidean plane $\mathbb{R}^2$ by setting $(x_1, x_2) \sim (y_1, y_2)$ if and only if $x_1 - y_1, x_2 - y_2 \in \mathbb{Z}$. Show that $\sim$ is an equivalence relation. What is the equivalence class of the origin $(0,0) \in \mathbb{R}^2$? What is common between a doughnut and the quotient space here?

Exercise A.2.8. Let us define a relation $\sim$ in the punctured Euclidean space $\mathbb{R}^3 \setminus \{(0,0,0)\}$ by setting $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$ if and only if $(x_1, x_2, x_3) = (ty_1, ty_2, ty_3)$ for some $t \in \mathbb{R}^+$. Prove that $\sim$ is an equivalence relation. What is common between a sphere and the quotient space here?

Orders

Definition A.2.9 (Partial order). A non-empty set X is *partially ordered* if there is a *partial order* $\leq$ on X. That is, $\leq$ is a relation from X to X, such that it is

1. *reflexive*: $x \leq x$ (for all $x \in X$);
2. *anti-symmetric*: if $x \leq y$ and $y \leq x$ then $x = y$ (for all $x, y \in X$);
3. *transitive*: if $x \leq y$ and $y \leq z$ then $x \leq z$ (for all $x, y, z \in X$).

We say that y *is greater than* x (or x *is less than* y), denoted by $x < y$, if $x \leq y$ and $x \neq y$.

Example. The set $\mathbb{R}$ of real numbers has the usual order $\leq$. Naturally, any of its non-empty subsets, e.g., $\mathbb{Z}^+ \subset \mathbb{R}$, inherits the order. The set $[-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$ has the order $\leq$ extended from $\mathbb{R}$, with conventions $-\infty \leq x$ and $x \leq +\infty$ for every $x \in [-\infty, +\infty]$.

Example. Let us order $X = \mathcal{P}(S)$ by inclusion. That is, for $A, B \subset S$, let $A \leq B$ if and only if $A \subset B$.

Example. Let X, Y be sets, where Y has a partial order $\leq$. We may introduce a new partial order for all functions $f, g : X \to Y$ by setting

$$f \leq g \quad \overset{\text{definition}}{\Longleftrightarrow} \quad \forall x \in X: \ f(x) \leq g(x).$$

This partial order is commonplace especially when $Y = \mathbb{R}$ or $Y = [-\infty, \infty]$.

Definition A.2.10 (Chains and total order). A non-empty subset $K \subset X$ is a *chain* if $x \leq y$ or $y \leq x$ for all $x, y \in K$. The partial order is *total* (or *linear*) if the whole set X is a chain.

Example. $[-\infty, +\infty]$ is a chain with the usual partial order. Thereby also its subsets are chains, e.g., $\mathbb{R}$ and $\mathbb{Z}^+$. If $\{A_j : \ j \in J\} \subset \mathcal{P}(S)$ is a chain then $A_j \subset A_k$ or $A_k \subset A_j$ for each $j, k \in J$. Moreover, $\mathcal{P}(S)$ is not a chain if S has more than one element.

Definition A.2.11 (Bounds). Let $\leq$ be a partial order on X. The sets of *upper and lower bounds* of $A \subset X$ are defined, respectively, by

$$\begin{aligned} \uparrow A &:= \{x \in X \mid \forall a \in A: \ a \leq x\}, \\ \downarrow A &:= \{x \in X \mid \forall a \in A: \ x \leq a\}. \end{aligned}$$

If $x \in A \cap \uparrow A$ then it is the *maximum of* A, denoted by $x = \max(A)$. If $x \in A \cap \downarrow A$ then it is the *minimum of* A, denoted by $x = \min(A)$. If $A \cap \uparrow \{z\} = \{z\}$ then the element $z \in A$ is called *maximal in* A. Similarly, if $A \cap \downarrow \{z\} = \{z\}$ then the element $z \in A$ is called *minimal in* A. If $\sup(A) := \min(\uparrow A) \in X$ exists, it is called the *supremum of* A, and if $\inf(A) := \max(\downarrow A) \in X$ exists, it is the *infimum of* A.

Remark A.2.12. Notations like

$$\sup_{k\geq 1} x_k = \sup_{k\in\mathbb{Z}^+} x_k = \sup\{x_k :\ k\in\mathbb{Z}^+\}$$

are quite common.

Example. The minimum in $\mathbb{Z}^+$ is 1, but there is no maximal element. For each $A\subset[-\infty,\infty]$, the infimum and the supremum exist.

Example. Let $X=\mathcal{P}(S)$. Then $\max(X)=S$ and $\min(X)=\emptyset$. If $\mathcal{A}\subset X$ then $\sup(\mathcal{A})=\bigcup\mathcal{A}$ and $\inf(\mathcal{A})=\bigcap\mathcal{A}$. For each $x\in S$, element $S\setminus\{x\}\in X$ is maximal in the subset $X\setminus\{S\}$.

Definition A.2.13 (lim sup and lim inf). Let $x_k\in X$ for each $k\in\mathbb{Z}^+$. If the following supremums and infimums exist, let

$$\limsup_{k\to\infty} x_k \;:=\; \inf\left\{\sup\{x_k:\ j\leq k\}\mid j\in\mathbb{Z}^+\right\},$$
$$\liminf_{k\to\infty} x_k \;:=\; \sup\left\{\inf\{x_k:\ j\leq k\}\mid j\in\mathbb{Z}^+\right\}.$$

Example. Let $E_k\in\mathcal{P}(X)$ for each $k\in\mathbb{Z}^+$. Then

$$\limsup_{k\to\infty} E_k = \bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty} E_k,$$
$$\liminf_{k\to\infty} E_k = \bigcup_{j=1}^{\infty}\bigcap_{k=j}^{\infty} E_k.$$

Exercise A.2.14. Let $A=\limsup_{k\to\infty} E_k$ and $B=\liminf_{k\to\infty} E_k$ as in the example above. Show that

$$\chi_A=\limsup_{k\to\infty}\chi_{E_k}\quad\text{and}\quad\chi_B=\liminf_{k\to\infty}\chi_{E_k},$$

where $\chi_E: X\to\mathbb{R}$ is the characteristic function of $E\subset X$.

A.3 Dominoes tumbling and transfinite induction

The principle of mathematical induction can be compared to a sequence of dominoes, falling over one after another when the first tumbles down. More precisely,

if $1\in S\subset\mathbb{Z}^+$ and $n\in S\Rightarrow n+1\in S$ for every $n\in\mathbb{Z}^+$, then $S=\mathbb{Z}^+$.

The Transfinite Induction Principle generalises this, working on any well-ordered set.

Definition A.3.1 (Well-ordered sets). A partially ordered set X is said to be *well ordered*, if $\min(A)$ exists whenever $\emptyset\neq A\subset X$.