Christoffel Revisited

Bas Fagginger Auer

September 1, 2009

Universiteit Utrecht

Master's Thesis
Mathematical Institute
Utrecht University
Supervisor: Prof. Dr. J. J. Duistermaat

This is an digitally edited photograph of Elwin Bruno Christoffel, the original of which was taken from the History of Mathematics archive of the School of Mathematics and Statistics of the University of St Andrews, Scotland, `http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Christoffel.html`.

Acknowledgements

Before we start with the actual thesis, I would very much like to thank a few people whose input and putting up with my ramblings have made it possible to finish this thesis in its current form. I am very much indebted to my supervisor, professor *Hans Duistermaat*, for introducing me to this subject and helping me overcome many mathematical difficulties (in particular showing me simpler and more elegant ways to prove a great number of results) during our fruitful and pleasant discussions. My girlfriend, *Hedwig van Driel*, for painstakingly proofreading this entire document, her care, stuffing me with food, our lovely evenings, and being firm with me when I was procrastinating; thank you.

Finally I would like to thank the numerous people whom I have bothered with my questions and unasked-for exposition of my results, in particular: *Mathijs Wintraecken*, *Matthijs van Dorp*, *Job Kuit*, *Albert-Jan Yzelman*, *Jan Jitse Venselaar*, and *Jaap Eldering*, as well as my parents for their continued support and interest.

Thank you all for your kind help,

Bas Fagginger Auer.

Abstract

This thesis discusses E. B. Christoffel's famous article from 1869 (an English translation is also included), generalises it to the more general setting of locally convex Hausdorff topological vector spaces over $\mathbb{R}$ or $\mathbb{C}$, and furthermore establishes a partial converse to the results of Christoffel in Banach spaces. Apart from this we rigorously discuss the aforementioned more general setting by introducing the concepts related to this setting from the ground up. This includes in particular a non-standard notion for the derivative and proofs of many useful statements, among which the open mapping theorem, closed graph theorem, fundamental theorem of integration, and the Taylor approximation theorem.

Contents

CHAPTER 1

INTRODUCTION

Welcome to this thesis, which is concerned with generalising E. B. Christoffel's article, [Chr1869], and developing the underlying theory of the setting in which this generalisation should take place. The level of the material contained in this thesis should be appropriate for any master student of Mathematics, as we develop the theory mostly from the ground up, relying only on results established in basic set theory and analysis on $\mathbb{R}$ and $\mathbb{C}$. However, a familiarity with topology, higher-dimensional analysis, and differential geometry will be very helpful for understanding the structure of this document, the included examples, and reasons for adopting certain definitions.

We are interested in Christoffel's article, because of its paramount importance for the development of differential geometry at the end of the 19th century. Through the introduction of the Christoffel symbols (in [Chr1869] denoted by $\begin{Bmatrix} ij \\ k \end{Bmatrix}$, but in differential geometry conventionally by Γ), the curvature tensor (in [Chr1869] denoted by $(ijkl)$ and now usually by R), and the means of covariant differentiation by using the Christoffel symbols, Christoffel provided very useful tools for the further development of differential geometry, as was undertaken by Gregorio Ricci-Curbastro and Tullio Levi-Civita (see [StAndrews]).

These developments in turn permitted Albert Einstein to formulate his theory of general relativity entirely in terms of differential geometry, which was a major step in the physical modeling of the effects of gravity and electromagnetism in celestial mechanics.

As differential geometry and general relativity are both still being practised by a great number of mathematicians and physicists today, this makes Christoffel's article highly influential and very interesting to further investigate. Even more so because of the geometrical way in which the Christoffel symbols are currently introduced in differential geometry (via an affine connection on a vector bundle, see [Ban2008]), which is not at all like the algebraic way in which they were used by Christoffel as tools to determine whether or not two given metrics could be transformed into one another via an appropriate coordinate transformation.

1.1 Notation

To ensure a concise treatment of the discussed material, we will strive to use the same symbols to denote the same type of objects. However, this is not always possible when a large number of objects is being discussed at the same time, so the following table is only meant to give an indication.

A, B, ...	Sets.
a, a', a_1, a_2, ...	Elements of the set A.
U, V, ...	Open subsets of A, B, ... respectively. See Definition (2.1.2).
$\mathcal{A}$, $\mathcal{B}$, ...	Collections of subsets of A, B, ... respectively.
f, g, ...	Functions between sets.
i, j, ...	Indices of objects.
k, l, ...	Elements of $\mathbb{N}$.
$\sim$	An equivalence relation.
α, β, ...	Scalars, usually values in either $\mathbb{R}$ or $\mathbb{C}$.

For collections of numbers we will employ the usual notation.

$\mathbb{N}$	The natural numbers: $1, 2, 3, \ldots$.
$\mathbb{N}_0$	The natural numbers together with zero: $0, 1, 2, 3, \ldots$.
$\hat{\mathbb{N}}$	The natural numbers extended with infinity and considered as a topological space: $1, 2, 3, \ldots, \infty$. See Example (2.3.2).
$\mathbb{Z}$	The integers: $\ldots, -2, -1, 0, 1, 2, 3, \ldots$.
$\mathbb{Q}$	The rationals.
$\mathbb{R}$	The real numbers.
$\mathbb{C}$	The complex numbers, identified with the plane $\mathbb{R}^2$.
$\mathbb{K}$	Refers to either $\mathbb{R}$ or $\mathbb{C}$.
$]\alpha, \beta[$	The open interval $\{\gamma \in \mathbb{R} \vert \alpha < \gamma < \beta\} \subseteq \mathbb{R}$.
$[\alpha, \beta]$	The closed interval $\{\gamma \in \mathbb{R} \vert \alpha \leq \gamma \leq \beta\} \subseteq \mathbb{R}$.
$]\alpha, \infty[$	The open interval $\{\gamma \in \mathbb{R} \vert \alpha < \gamma\} \subseteq \mathbb{R}$.
$]-\infty, \alpha[$	The open interval $\{\gamma \in \mathbb{R} \vert \alpha > \gamma\} \subseteq \mathbb{R}$.
$]-\infty, \infty[$	The real line $\mathbb{R}$.

As well as the following notation for set operations.

$\emptyset$	The *empty set*.
$A \setminus B$	The *complement* of the set B in A, $\{a \in A \vert a \notin B\}$.
$\bigcup \mathcal{A}$	The *union* of all sets in $\mathcal{A}$, $\{a \vert \exists A \in \mathcal{A} : a \in A\}$.
$\bigcup_{i \in I} A_i$	Defined as $\bigcup \{A_i \vert i \in I\}$.
$A_1 \cup \ldots \cup A_k$	Defined as $\bigcup \{A_1, \ldots, A_k\}$.
$\bigcap \mathcal{A}$	The *intersection* of all sets in $\mathcal{A}$, $\{a \vert \forall A \in \mathcal{A} : a \in A\}$.
$\bigcap_{i \in I} A_i$	Defined as $\bigcap \{A_i \vert i \in I\}$.
$A_1 \cap \ldots \cap A_k$	Defined as $\bigcap \{A_1, \ldots, A_k\}$.
$\coprod \mathcal{A}$	The *disjoint union* of all sets in $\mathcal{A}$, the set $\{(A, a) \vert A \in \mathcal{A}, a \in A\}$.
$\coprod_{i \in I} A_i$	Defined as $\{(i, a) \vert i \in I, a \in A_i\}$.
$\prod \mathcal{A}$	The *product* of all sets in $\mathcal{A}$, the set $\{g : \mathcal{A} \to \bigcup \mathcal{A} \vert \forall A \in \mathcal{A} : g(A) \in A\}$.
$\prod_{i \in I} A_i$	Defined as $\{g : I \to \bigcup_{i \in I} A_i \vert \forall i \in I : g(i) \in A_i\}$.
$A_1 \times \ldots \times A_k$	Defined as $\{(a_1, \ldots, a_k) \vert a_1 \in A_1, \ldots, a_k \in A_k\}$.

Together with their usual identifications. [1] [2] [3]

We will also use the following symbols.

$\operatorname{dom} f$	The *domain* of a function, for $f : A \to B$, $\operatorname{dom} f := A$.
$\operatorname{im} f$	The *image* of a function, for $f : A \to B$, $\operatorname{im} f := f(A) := \{b \in B \vert \exists a \in A : f(a) = b\} \subseteq B$.
graph	The *graph* of a function $f : A \to B$, $\operatorname{graph} f := \{(a, b) \in A \times B \vert f(a) = b\} \subseteq A \times B$.
$f^{-1}(\cdot)$	The *pre-image* of a set $C \subseteq B$ under a function $f : A \to B$, $f^{-1}(C) := \{a \in A \vert f(a) \in C\} \subseteq A$.
id_A	The *identity map*, for a given set A, $\mathrm{id}_A := A \to A : a \mapsto a$.
sgn	The *sign* of a number, $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, +1\}$ where $\operatorname{sgn}(\alpha) := \begin{cases} -1 & \alpha < 0 \\ 0 & \alpha = 0 \\ +1 & \alpha > 0 \end{cases}$.
$\operatorname{Re}, \operatorname{Im}$	The *real* and *imaginary* parts of a complex number, for $z = (x, y) = x + i\,y \in \mathbb{C} \simeq \mathbb{R}^2$, $\operatorname{Re}(z) := x$, $\operatorname{Im}(z) := y$.
$\mathcal{P}(A)$	The *collection of subsets* or *powerset* of a set A, $\mathcal{P}(A) := \{B \vert B \subseteq A\}$.
$\operatorname{int}(A)$	*Interior* of a set A, Definition (2.1.2).
$\overline{A}$	*Closure* of a set A, Definition (2.1.2).
$\mathcal{T}(A)$	The *topology generated by* A, Definition (2.1.6).
S^k	Group of permutations of $\{1, \ldots, k\}$, Example (3.1.4).
Abc	Absorbent, balanced, and convex, Definition (4.3.4).
$D^k_a f$	The k-th *derivative* of a function f at a, Definition (5.1.1) and Definition (5.1.9).
$C^k(U, B)$	The set of all k*-times continuously differentiable functions* from an open set $U \subseteq A$ to B, Definition (5.1.9).
$\int_\alpha^\beta f$	The *integral* of a function f over the interval $[\alpha, \beta]$, see Definition (5.3.2).
$L(A, B)$	Space of all continuous linear maps between Banach spaces A and B, see Definition (5.5.4).
$e^{\cdot f}$	The *flow* of f, Theorem (5.5.10).

Where our notation truly differs from what is usual, is by denoting properties of objects. Because later objects, in particular topological vector spaces, can have a large number of different properties, these properties are cumbersome to fully write out in words. Therefore we will employ a shorthand in the form of coloured icons, each of which denotes an object type or property.

For algebraic objects, we furthermore employ $B \leq A$ to indicate that $B \subseteq A$ and that B is an algebraic object of the same type as A, with regard to the restrictions of the algebraic operators (i.e. addition, multiplication, ...) from A to B.

[1] We identify $\coprod_{i \in I} A_i$ with $\coprod\{A_i \vert i \in I\}$ via $(i, a) \leftrightarrow (A_i, a)$.
[2] We identify $\prod_{i \in I} A_i$ with $\prod\{A_i \vert i \in I\}$ via $(A_i \mapsto g(A_i)) \leftrightarrow (i \mapsto g(A_i))$.
[3] We identify $A_1 \times \ldots \times A_k$ with $\prod_{i \in \{1, \ldots, k\}} A_i$ via $(a_1, \ldots, a_k) \leftrightarrow (i \mapsto a_i)$.

T	Topological space, Definition (2.1.1).
c	Continuity, Definition (2.1.11).
T0-T6	Separation class, Definition (2.2.1).
1C	First countability, Definition (2.3.3).
Cpt	Compactness, Definition (2.4.2).
d(·,·)	Metric space, Definition (2.5.1).
G	Group, Definition (3.1.1).
R	Ring, Definition (3.2.1).
F	Field, Definition (3.2.6).
M	Module, Definition (3.3.1).
l	Linearity, Definition (3.3.3).
Vs	Vector space, Definition (3.3.21).
T_G	Topological group, Definition (4.1.1).
T_R	Topological ring, Definition (4.1.2).
T_M	Topological module, Definition (4.1.3).
‖·‖	Normed module, Definition (4.2.3).
T_F	Topological field, Definition (4.3.1).
T_{Vs}	Topological vector space, Definition (4.3.3).
FS	F-space, Definition (4.4.1).
LC	Local convexity, Definition (4.5.1).
UC	Uniform completeness, Definition (4.5.9).
d	Differentiability, Definition (5.1.1).
Fr	Fréchet space, Definition (5.4.1).
Ba	Banach space, Definition (5.5.1).

It should be noted that all icons representing topological properties are coloured green, the icons representing algebraical properties blue, and the icons representing properties depending on both topology and algebra green-blue. Furthermore, these icons permit us to easily talk about maps preserving a certain property: instead of a homeomorphism between two topological spaces we can talk about a T-isomorphism, instead of a homomorphism we can talk about a G-morphism, etc.. This ensures that we do not have to create names for different types of maps preserving different properties, but call all of them 'morphisms' with respect to a certain icon. [4]

1.2 Overview

We will start in Chapter 2 by discussing the basic concepts of topologies, continuity, and metric spaces. Of particular interest here are: chain rule for limits (Lemma (2.1.10)), initial and final topologies (Definition (2.1.18) and Definition (2.1.23)), operations preserving continuity (Theorem (2.1.28)), separation axioms (Definition (2.2.1) and Theorem (2.2.5)), graphs of continuous functions being closed (Lemma (2.2.7)), notion of compactness (Definition (2.4.2)), fixed point theorem (Theorem (2.5.21)), and uniform continuity theorem (Theorem (2.5.23)).

We then continue to discuss the basics of group theory and algebra (in the form of rings, fields, and modules) in Chapter 3. Interesting notions here are: factorisation lemma for G-morphisms (Lemma (3.1.9)), group action (Definition

[4]Inspired by category theory.

(3.1.10)), solving equations in a ring (Lemma (3.2.3)), factorisation lemma for modules (Lemma (3.3.10)), duality theorem (Theorem (3.3.19)), and the first version of the Hahn-Banach theorem (Theorem (3.3.22) and Example (3.3.23)).

After Chapter 3 we combine topology and algebra in Chapter 4 where we introduce algebraical objects of which all algebraical manipulations should be continuous (topological rings, fields, and modules). We discuss (semi)normed spaces, F-space, and the notion of local convexity. Of particular interest are the fact that translations and non-zero scalings of open sets are open (Lemma (4.1.5)), the comparison between topologies of topological modules (Lemma (4.1.6)), separation of topological modules (Lemma (4.1.7)), notion of a seminorm (Definition (4.2.3)), the second, more definitive form of the Hahn-Banach theorem (Theorem (4.2.6)), topological vector spaces (Definition (4.3.3)), absorbent balanced and convex subsets (Definition (4.3.4), Lemma (4.3.5), and Lemma (4.3.6)), counterexamples emphasising that we need to be careful in this general setting (Example (4.3.7), Example (4.3.8), and Example (4.3.9)), open mapping and closed graph theorems (Theorem (4.4.3), Theorem (4.4.4), Theorem (4.4.5), and Corollary (4.4.6)), notion of local convexity (Section 4.5 entirely), the final form of the Hahn-Banach theorem (Theorem (4.5.14)), and comparison of a space and the topological dual of its topological dual (Theorem (4.5.15)).

Now that we have established our definitive setting (in the form of topological vector spaces) for the generalisation of [Chr1869], we start doing analysis on topological vector spaces in Chapter 5. Here we discuss a slightly different generalised notion of differentiability (when comparing with the Fréchet and Gâteaux derivatives that are usually employed) in Definition (5.1.1). We define partial derivatives in Definition (5.1.12), higher order derivatives in Definition (5.1.9)), and verify this different notion to be a proper generalisation in Theorem (5.1.5), Corollary (5.1.7), and Corollary (5.5.7). This derivative furthermore has the usually desired properties as expressed in Theorem (5.1.8) (most notably the chain rule and local constantness), Lemma (5.1.11), Lemma (5.1.13) (sum rule), and Theorem (5.1.16) (symmetry for higher order derivatives). We then investigate differentiability of families of multilinear maps, Definition (5.2.2), in Lemma (5.2.1) and establish a very useful product rule as Equation (5.4) in Theorem (5.2.4), and finally a condition for differentiability of families of linear inverses in Theorem (5.2.6). After differentiation we consider integration in Definition (5.3.2) for which we prove the usual properties in Theorem (5.3.5) and the fundamental theorem of integration in Theorem (5.3.8). These results are then used to show the Taylor approximation theorem in Theorem (5.3.11) and Corollary (5.3.12). We conclude by discussing Fréchet spaces (Definition (5.4.1)) and Banach spaces (Definition (5.5.1)) and showing that while the inverse function theorem (Theorem (5.5.8)), implicit function theorem (Theorem (5.5.9)), and existence of solutions for ordinary differential equations (Theorem (5.5.10)) are all true for Banach spaces, they do not hold for Fréchet spaces, as shown in Example (5.4.3) and Example (5.4.4).

After this we have derived all necessary theory, and in Chapter 6 we generalise [Chr1869] to Theorem (6.2.1), Theorem (6.2.2), and Theorem (6.2.3) and establish a partial converse in Theorem (6.6.1).

We conclude the thesis with an English translation of the originally German article [Chr1869] in Chapter 8 and a conclusion in Chapter 7.

Chapter 2

Topology

Topology is the study of qualitative geometry: we provide a given set A that has a geometrical interpretation with a notion of what it means to 'be near' or 'in a neighbourhood of' a point in A[1] by selecting a certain collection of subsets of A that are all interpreted as 'neighbourhoods'. These subsets are called *open* sets in A and they give a surprising amount of information about the geometrical properties of A (see for example [Mun2000]).

2.1 Topological spaces

Definition 2.1.1: Topology ($\boxed{\mathrm{T}}$)
Let A be a set.

Then a *topology on* A is a collection $\mathcal{A} \subseteq \mathcal{P}(A)$ of subsets of A such that

- $\forall \mathcal{U} \subseteq \mathcal{A} : \bigcup \mathcal{U} \in \mathcal{A}$,
- $\forall U_1, U_2 \in \mathcal{A} : U_1 \cap U_2 \in \mathcal{A}$,
- $\emptyset, A \in \mathcal{A}$.

A *topological space* A is a set A together with a topology $\mathcal{A}$ on A, we will write $A\ \boxed{\mathrm{T}}$ to indicate that A is a topological space.

Elements $a \in A$ of a topological space are commonly called *points* to emphasise their geometrical interpretation.

Definition 2.1.2
Let $A\ \boxed{\mathrm{T}}$. Denote A's topology by $\mathcal{A}$.

A subset $U \subseteq A$ is called *open* if $U \in \mathcal{A}$ and *closed* if $A \setminus U \in \mathcal{A}$. For any subset $B \subseteq A$ we define the *closure*, and *interior* as

$$\overline{B} := \bigcap \{C \subseteq A \mid C \text{ closed}\,, B \subseteq C\}, \qquad \operatorname{int}(B) := \bigcup \{U \subseteq A \mid U \text{ open}\,, U \subseteq B\}$$

respectively.

Let $a \in A$ be any point, then we call a subset $B \subseteq A$ *a neighbourhood of a in* A if there exists an open set $U \subseteq A$ for which $a \in U \subseteq B$.

[1] But exactly *how* near is left unspecified.

Note that for any point $a \in U$ of an open set $U \subseteq A$, U is an (open) neighbourhood of a in A.

Lemma 2.1.3
Let A T.

Then for any subset $B \subseteq A$,

- $\overline{B}$ is the smallest closed set containing B,
- $\mathrm{int}(B)$ is the largest open set contained in B,
- $a \in \overline{B}$ if and only if for all open neighbourhoods U of a in A we have $U \cap B \neq \emptyset$,
- $a \in \mathrm{int}(B)$ if and only if there exists an open neighbourhood U of a in A such that $a \in U \subseteq B$.

Proof. $\overline{B}$ is nonempty if B is nonempty since $A \supseteq B$ is closed. As arbitrary unions of open sets are open and closed sets are complements of open sets, arbitrary intersections of closed sets are closed, so $\overline{B}$ is closed. From the definition $\overline{B}$ is clearly the smallest closed set containing B. The second item is proven in the same way.

Let $a \in A$.

Suppose there exists an open neighbourhood U of a in A with $U \cap B = \emptyset$, then $B \subseteq A \setminus U$ which is closed, so $\overline{B} \subseteq A \setminus U$. As $a \in U$, $a \notin A \setminus U \supseteq \overline{B}$, $a \notin \overline{B}$. Suppose conversely that $a \notin \overline{B}$, then $a \in A \setminus \overline{B}$ which is open ($\overline{B}$ is closed), so $A \setminus \overline{B}$ is an open neighbourhood of a in A and since $B \subseteq \overline{B}$, we have $(A \setminus \overline{B}) \cap B = \emptyset$. This shows the third item.

Suppose $a \in \mathrm{int}(B)$, then $\mathrm{int}(B)$ is an open neighbourhood of a in A and $a \in \mathrm{int}(B) \subseteq B$. If conversely there exists an open neighbourhood U of a in A such that $a \in U \subseteq B$, then $a \in U \subseteq \mathrm{int}(B)$ by definition since U is an open set contained in B. This shows the fourth item. $\square$

Definition 2.1.4: Neighbourhood basis
Let A T and $a \in A$.

Then we call a collection $\mathcal{A} \subseteq \mathcal{P}(A)$ a *basis of neighbourhoods of a in A* if for all $U \in \mathcal{A}$, U is a neighbourhood of a in A and for all open neighbourhoods U_1 of a in A, there exists some $U \in \mathcal{A}$ such that $U \subseteq U_1$.

Definition 2.1.5: Topological basis
Let A be a set.

Then a *topological basis on* A is a collection $\mathcal{A} \subseteq \mathcal{P}(A)$ such that

- $A = \bigcup \mathcal{A}$,
- $\forall U_1, U_2 \in \mathcal{A} : \exists U_3 \in \mathcal{A} : U_3 \subseteq U_1 \cap U_2$.

Note that any topology is a topological basis, but that the converse is not necessarily true.

Definition 2.1.6: Generated topology
Let A be a set and $\mathcal{A} \subseteq \mathcal{P}(A)$ a collection of subsets.

Then *the topology generated by* $\mathcal{A}$ is defined to be the intersection of all topologies on A containing $\mathcal{A}$:

$$\mathcal{T}(\mathcal{A}) := \bigcap \{\mathcal{A}_1 \subseteq \mathcal{P}(A) | \mathcal{A} \subseteq \mathcal{A}_1 \text{ and } \mathcal{A}_1 \text{ is a topology on } A\}.$$

Lemma 2.1.7: Properties of $\mathcal{T}(\cdot)$
Let A be a set and $\mathcal{A} \subseteq \mathcal{P}(A)$.

Then

- $\mathcal{T}(\mathcal{A})$ is the unique smallest topology on A containing $\mathcal{A}$,
- if $\mathcal{A}$ is a topology, then $\mathcal{T}(\mathcal{A}) = \mathcal{A}$,
- if $\mathcal{A}$ is a topological basis, then $U \in \mathcal{T}(\mathcal{A})$ if and only if for all $a \in U$ there exists a $U_a \in \mathcal{A}$ such that $a \in U_a \subseteq U$.

Proof. The first item is direct from the definition of $\mathcal{T}(\mathcal{A})$ as the intersection of all topologies (which is directly verified to again be a topology) containing $\mathcal{A}$, by it being the smallest, it is also unique.

The second item follows directly from the first item, since in this case $\mathcal{A}$ itself is the smallest topology containing $\mathcal{A}$.

For the third item, suppose $\mathcal{A}$ is a topological basis and let

$$\mathcal{A}_1 := \{U \subseteq A | \forall a \in U : \exists U_a \in \mathcal{A} : a \in U_a \subseteq U\}.$$

Then $\emptyset \in \mathcal{A}_1$ vacuously and $A \in \mathcal{A}_1$ because $A = \bigcup \mathcal{A}$. Clearly for all $\mathcal{U} \subseteq \mathcal{A}_1$ we have $\bigcup \mathcal{U} \in \mathcal{A}_1$. For $U_1, U_2 \in \mathcal{A}_1$ we have $U_1 \cap U_2 \in \mathcal{A}_2$, because for all $a \in U_1 \cap U_2$ there exist $U_3, U_4 \in \mathcal{A}$ such that $a \in U_3 \subseteq U_1$ and $a \in U_4 \subseteq U_2$, now as $\mathcal{A}$ is a basis there exists a $U_5 \in \mathcal{A}$ with $a \in U_5 \subseteq U_3 \cap U_4 \subseteq U_1 \cap U_2$. Therefore $\mathcal{A}_1$ is a topology and for any $U \in \mathcal{A}$ and $a \in U$ we have $a \in U \subseteq U$, so $\mathcal{A} \subseteq \mathcal{A}_1$. Therefore $\mathcal{T}(\mathcal{A}) \subseteq \mathcal{A}_1$. Now let $\mathcal{A}_2$ be any topology containing $\mathcal{A}$. Let $U \in \mathcal{A}_1$, then for all $a \in U$ there exists a $U_a \in \mathcal{A}$ such that $a \in U_a \subseteq U$. As all $U_a \in \mathcal{A} \subseteq \mathcal{A}_2$, we have $U = \bigcup_{a \in U} U_a \in \mathcal{A}_2$. Since this is true for all $U \in \mathcal{A}_1$, $\mathcal{A}_1 \subseteq \mathcal{A}_2$. Because this is true for all topologies $\mathcal{A}_2$ containing $\mathcal{A}$ we have $\mathcal{A}_1 \subseteq \mathcal{T}(\mathcal{A})$. Therefore $\mathcal{A}_1 = \mathcal{T}(\mathcal{A})$. □

Lemma 2.1.8
Let A be a set and $\mathcal{A}_1$, $\mathcal{A}_2$ topologies on A. Let $\mathcal{B}_1$, $\mathcal{B}_2$ be any topological bases on A satisfying $\mathcal{A}_1 = \mathcal{T}(\mathcal{B}_1)$ and $\mathcal{A}_2 = \mathcal{T}(\mathcal{B}_2)$.

Then $\mathcal{A}_1 \subseteq \mathcal{A}_2$ if and only if for all $U_1 \in \mathcal{B}_1$ and $a \in U_1$ there exists a $U_2 \in \mathcal{B}_2$ such that $a \in U_2 \subseteq U_1$.

Proof. Suppose $\mathcal{A}_1 \subseteq \mathcal{A}_2$. Let $U_1 \in \mathcal{B}_1$ be arbitrary and $a \in U_1$, then because $U_1 \in \mathcal{B}_1 \subseteq \mathcal{T}(\mathcal{B}_1) = \mathcal{A}_1 \subseteq \mathcal{A}_2$ we have $a \in U_2 \subseteq U_1$ for some $U_2 \in \mathcal{B}_2$ by Lemma (2.1.7).

Suppose conversely that for all $U_1 \in \mathcal{B}_1$ and $a \in U_1$ there exists a $U_2 \in \mathcal{B}_2$ with $a \in U_2 \subseteq U_1$. Let $U_1 \in \mathcal{A}_1$, then by Lemma (2.1.7) ($\mathcal{B}_1$ is a topological basis), for all $a \in U_1$ there exists a $U_a \in \mathcal{B}_1$ such that $a \in U_a \subseteq U_1$. Now by our assumption, for all such U_a, $a \in U_a$ there exists a $U_a' \in \mathcal{B}_2$ such that $a \in U_a' \subseteq U_a \subseteq U_1$. Therefore for all $a \in U_1$ there exists a $U_a' \in \mathcal{B}_2$ such that $a \in U_a' \subseteq U_1$, so (again by Lemma (2.1.7)) $U_1 \in \mathcal{A}_2$. Therefore $\mathcal{A}_1 \subseteq \mathcal{A}_2$. □

Definition 2.1.9: Limit
Let A, B T, $f : A \to B$, $a \in A$, and $b \in B$.

Then we say that f *has limit* b *at* a, denoted by

$$\lim_{x \to a} f(x) = b,$$

if for all neighbourhoods V of b in B, $f^{-1}(V)$ is a neighbourhood of a in A.

If it is not true that $\lim_{x\to a} f(x) = b$ we write

$$\lim_{x\to a} f(x) \neq b.$$

Note that $\lim_{x\to a} f(x) = b$ if and only if for all open $V \subseteq B$, $b \in V$ there exists an open $U \subseteq A$, $a \in U$ with $f(U) \subseteq V$.

Lemma 2.1.10: Chain rule for limits

Let A, B, C $\boxed{\mathrm{T}}$ and $f : A \to B$, $g : B \to C$ maps.

Suppose for $a \in A$, $b \in B$, and $c \in C$ we have

$$\lim_{x\to a} f(x) = b, \qquad \lim_{y\to b} g(y) = c,$$

then

$$\lim_{x\to a} (g \circ f)(x) = c.$$

Proof. Suppose the conditions for the lemma are met. Let W be a neighbourhood of c in C. Then because $\lim_{y\to b} g(y) = c$, $g^{-1}(W)$ is a neighbourhood of b in B. Hence, by $\lim_{x\to a} f(x) = b$, $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$ is a neighbourhood of a in A. So for any neighbourhood W of c in C, $(g \circ f)^{-1}(W)$ is a neighbourhood of a in A, therefore $\lim_{x\to a}(g \circ f)(x) = c$. $\square$

Definition 2.1.11: Continuous maps (Ⓒ)

Let A, B $\boxed{\mathrm{T}}$ and $a \in A$.

Then a map $f : A \to B$ is called *continuous at* a (denoted by f Ⓒ a) if

$$\lim_{x\to a} f(x) = f(a).$$

We call f *continuous* (f Ⓒ) if f is continuous at a for all $a \in A$.

Definition 2.1.12: Almost continuous maps

Let A, B $\boxed{\mathrm{T}}$, and $a \in A$.

Then a map $f : A \to B$ is called *almost continuous at* a if for each neighbourhood V of $f(a)$ in B, $\overline{f^{-1}(V)}$ is a neighbourhood of a in A.

Definition 2.1.13: Open and closed maps

Let A, B $\boxed{\mathrm{T}}$.

Then a map $f : A \to B$ is called *open* (resp. *closed*) if for any $U \subseteq A$ open (resp. closed), $f(U) \subseteq B$ is open (resp. closed).

A map $f : A \to B$ is called *almost open* if for any $U \subseteq A$ open, $f(U) \subseteq \operatorname{int}(\overline{f(U)})$.

Note that a map is almost open if and only if for each $a \in A$ and each open neighbourhood U of a in A, $\overline{f(U)}$ is a neighbourhood of $f(a)$. For if this holds and $U \subseteq A$ is open and arbitrary, there exists for each $a \in U$ an open neighbourhood V_a of $f(a)$ in B such that $V_a \subseteq \overline{f(U)}$ and therefore $f(U) = \bigcup_{a\in U}\{f(a)\} \subseteq \bigcup_{a\in U} V_a \subseteq \operatorname{int}(\overline{f(U)})$, as all V_a are open and contained in $\overline{f(U)}$. Conversely, let $a \in A$ and U be an open neighbourhood of a in A, then $f(a) \in f(U) \subseteq \operatorname{int}(\overline{f(U)}) \subseteq \overline{f(U)}$, so $\overline{f(U)}$ is a neighbourhood of $f(a)$ in B. This makes both characterisations equivalent.

Lemma 2.1.14
Let A, B $\boxed{\mathrm{T}}$ and $f : A \to B$ a map.

Then f Ⓒ if and only if for all $V \subseteq B$ open we have that $f^{-1}(V)$ is open in A, and if and only if for all $V \subseteq B$ closed we have that $f^{-1}(V)$ is closed in A.

If the topology of B is generated by a topological basis $\mathcal{B}$, f Ⓒ if and only if for all $V \in \mathcal{B}$ we have that $f^{-1}(V)$ is open in A.

Proof. Suppose f Ⓒ and let $V \subseteq B$ be open. Let $a \in f^{-1}(V)$, then because f Ⓒ by assumption, f Ⓒ a and hence $\lim_{x\to a} f(x) = f(a) \in V$. By definition of limit and the fact that V is an open neighbourhood of $f(a)$ in B there exists an open neighbourhood U_a of a in A such that $f(U_a) \subseteq V$, so $U_a \subseteq f^{-1}(V)$. Because of this $f^{-1}(V) = \bigcup_{a\in f^{-1}(V)}\{a\} \subseteq \bigcup_{a\in f^{-1}(V)} U_a \subseteq f^{-1}(V)$ as $a \in U_a \subseteq V$ for all $a \in f^{-1}(V)$. Therefore $f^{-1}(V) = \bigcup_{a\in f^{-1}(V)} U_a$ is open in A as all U_a are open.

Suppose conversely that $f^{-1}(V)$ is open in A for all $V \subseteq B$ open. Let $a \in A$ be arbitrary and V an open neighbourhood of $f(a)$ in B. Then V is open, so by assumption $f^{-1}(V)$ is open in A and as $f(a) \in V$ we have that $a \in f^{-1}(V)$. This means that for all open neighbourhoods V of $f(a)$ in B, $f^{-1}(V)$ is an open neighbourhood of a in A and $f(f^{-1}(V)) \subseteq V$, so $\lim_{x\to a} f(x) = f(a)$ and f Ⓒ a. As this is true for all $a \in A$, f Ⓒ.

The same is true for closed sets, since $B \subseteq V$ is closed iff $V \setminus B$ is open and $f^{-1}(B \setminus V) = A \setminus f^{-1}(V)$ is open iff $f^{-1}(V)$ is closed.

For the second point it is sufficient to note that a set $V \subseteq B$ is open if and only if for all $b \in V$ there exists a $V_b \in \mathcal{B}$ such that $b \in V_b \subseteq V$ (by Lemma (2.1.7)). Therefore $f^{-1}(V) = f^{-1}(\bigcup_{b\in V} V_b) = \bigcup_{b\in V} f^{-1}(V_b)$ which is open as a union of open sets. $\square$

Example 2.1.15: Continuity is a local property
Consider $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) := \begin{cases} x^2 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases},$$

then f Ⓒ 0 ($\lim_{x\to 0} f(x) = 0 = f(0)$), but for all $x \in \mathbb{R} \setminus \{0\}$ we have that not f Ⓒ x.

Hence this function is continuous in just a single point.

Lemma 2.1.16
Let A, B $\boxed{\mathrm{T}}$ and $f : A \to B$ Ⓒ.

Then for any subset $C \subseteq A$, we have

$$f(\overline{C}) \subseteq \overline{f(C)}.$$

Proof. Let $b \in f(\overline{C})$, then there exists an $a \in \overline{C}$ such that $f(a) = b$. Let V be any open neighbourhood of b in B. Then, as f Ⓒ, $\lim_{x\to a} f(x) = f(a) = b$, there exists an open neighbourhood U of a in A such that $f(U) \subseteq V$. As $a \in \overline{C}$, by Lemma (2.1.3) there exists an $a_1 \in U \cap C$. Hence $f(a_1) \in f(U \cap C) \subseteq V \cap f(C)$. Therefore $V \cap f(C) \neq \emptyset$ for any open neighbourhood V of b in B, so $b \in \overline{f(C)}$ by Lemma (2.1.3). $\square$

That we do not necessarily have equality of $f(\overline{C})$ and $\overline{f(C)}$ is shown in the following example.

Example 2.1.17

Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) := e^{-x}$ which is Ⓒ.

Now $f(]0,\infty[) =]0,1[$, but $f(\overline{]0,\infty[}) = f([0,\infty[) =]0,1] \subsetneq [0,1] = \overline{f(]0,\infty[)}$. So in this case

$$f(\overline{]0,\infty[}) \subsetneq \overline{f(]0,\infty[)}.$$

Definition 2.1.18: Initial topology

Let A be a set, $\{f_i : A \to B_i | i \in I\}$ a collection of maps with B_i $\boxed{\text{T}}$ for all $i \in I$.

Then we define the *initial topology on A with respect to* $\{f_i : A \to B_i | i \in I\}$ as

$$\mathcal{T}(\{f_i^{-1}(V) \subseteq A \mid V \subseteq B_i \text{ open}, i \in I\}).$$

Lemma 2.1.19: Properties of the initial topology

Let A be a set, $\{f_i : A \to B_i | i \in I\}$ a collection of maps with B_i $\boxed{\text{T}}$ for all $i \in I$.

Let the topology on A be the initial topology with respect to this collection.

- The initial topology is the smallest topology for which f_i Ⓒ for all $i \in I$.
- Suppose that for all $i \in I$, the topology on B_i is generated by the topological basis $\mathcal{B}_i$. Then the collection

  $$\mathcal{A} := \{f_{i_1}^{-1}(V_1) \cap \ldots \cap f_{i_k}^{-1}(V_k) \subseteq A \\ |k \in \mathbb{N}, i_1, \ldots, i_k \in I \text{ distinct}, V_1 \in \mathcal{B}_{i_1}, \ldots, V_k \in \mathcal{B}_{i_k}\}$$

  forms a topological basis for the initial topology.
- If furthermore all the B_i are equal to B $\boxed{\text{T}}$ with topology generated by the topological basis $\mathcal{B}$, then

  $$\mathcal{A}' := \{f_{i_1}^{-1}(V) \cap \ldots \cap f_{i_k}^{-1}(V) \subseteq A \\ |k \in \mathbb{N}, i_1, \ldots, i_k \in I \text{ distinct}, V \in \mathcal{B}\}$$

  forms a basis for the initial topology.
- For any C $\boxed{\text{T}}$ and $g : C \to A$ we have that g Ⓒ if and only if for all $i \in I$, $f_i \circ g : C \to B_i$ Ⓒ. The initial topology is the unique topology on A with this property.

Proof.
- The first item is direct from Lemma (2.1.14) together with the definition of the initial topology and Lemma (2.1.7).
- Now we consider the second item. First of all note that for any $i \in I$, $B_i = \bigcup \mathcal{B}_i$, so $A \supseteq \bigcup \mathcal{A} \supseteq \bigcup_{V \in \mathcal{B}_i} f_i^{-1}(V) = f_i^{-1}(\bigcup \mathcal{B}_i) = f_i^{-1}(B_i) = A$ and therefore $A = \bigcup \mathcal{A}$. Now let $U_1 = f_{i_1}^{-1}(V_1) \cap \ldots \cap f_{i_k}^{-1}(V_k), U_2 = f_{j_1}^{-1}(W_1) \cap \ldots \cap f_{j_l}^{-1}(W_l) \in \mathcal{A}$. Let $a \in U_1 \cap U_2$ be arbitrary, we are going to construct a $U_3 \in \mathcal{A}$ such that $a \in U_3 \subseteq U_1 \cap U_2$. Suppose that any of the $i_1, \ldots, i_k$ equals any of the $j_1, \ldots, j_l$, say $i_1 = j_1$. Then $a \in f_{i_1}^{-1}(V_1) \cap f_{j_1}^{-1}(W_1) = f_{i_1}^{-1}(V_1 \cap W_1)$, so $f_{i_1}(a) \in V_1 \cap W_1$. Since $V_1, W_1 \in \mathcal{B}_{i_1}$ which is a basis, there exists a $V' \in \mathcal{B}_{i_1}$ such that $f_{i_1}(a) \in V' \subseteq V_1 \cap W_1$, so $a \in f_{i_1}^{-1}(V')$. Now replace $f_{i_1}^{-1}(V_1) \cap f_{j_1}^{-1}(W_1)$ in $U_1 \cap U_2$ by $f_{i_1}^{-1}(V')$. Continue this way until all of (the finite number of indices) $i_1, \ldots, i_k, j_1, \ldots, j_l$ are distinct to obtain U_3. By construction

$a \in U_3$ and $U_3 \subseteq U_1 \cap U_2$ and because all indices are distinct $U_3 \in \mathcal{A}$, so $a \in U_3 \subseteq U_1 \cap U_2$ for $U_3 \in \mathcal{A}$. Since this can be done for all $U_1, U_2 \in \mathcal{A}$ and $a \in U_1 \cap U_2$, and $A = \bigcup \mathcal{A}$, $\mathcal{A}$ is a topological basis on A. Note that $\mathcal{A}$ is contained in the initial topology on A (all $f_{i_1}^{-1}(V_1) \cap \ldots \cap f_{i_k}^{-1}(V_k)$ are open by Lemma (2.1.14): f_{i_j} Ⓒ, V_j open), and that the collection generating the initial topology is contained in $\mathcal{A}$. Therefore $\mathcal{A}$ is a topological basis generating the initial topology.

- Now if all B_i are equal to B and all bases $\mathcal{B}_i$ are equal to $\mathcal{B}$, then we can for any $i_1, \ldots, i_k \in I$ and $V_{i_1}, \ldots, V_{i_k} \in \mathcal{B}$ consider $V_{i_1} \cap \ldots \cap V_{i_k}$ which is open in B and therefore there exists a $V \in \mathcal{B}$ such that $V \subseteq V_{i_1} \cap \ldots \cap V_{i_k}$. This shows that $\mathcal{A}'$ is at least as large as $\mathcal{A}$. On the other hand, $\mathcal{A}$ is clearly at least as large as $\mathcal{A}'$ (pick $V_{i_1} = \ldots = V_{i_k} = V$), hence $\mathcal{A} = \mathcal{A}'$ in this case and therefore $\mathcal{A}'$ forms a basis for the initial topology.

- For the final item, let C $\boxed{\mathrm{T}}$ and $g : C \to A$ be given. Suppose g Ⓒ, then as all f_i Ⓒ in the initial topology we have (Lemma (2.1.10)) that all $f_i \circ g$ Ⓒ.

  Suppose conversely that for all $i \in I$, $f_i \circ g$ Ⓒ. Let $f_{i_1}^{-1}(V_1) \cap \ldots \cap f_{i_k}^{-1}(V_k)$ be any element from the basis generating the initial topology. Then $g^{-1}(f_{i_1}^{-1}(V_1) \cap \ldots \cap f_{i_k}^{-1}(V_k)) = (f_{i_1} \circ g)^{-1}(V_1) \cap \ldots \cap (f_{i_k} \circ g)^{-1}(V_k) \subseteq C$ which is open, because all $f_{i_1} \circ g, \ldots, f_{i_k} \circ g$ Ⓒ and $V_1 \subseteq B_{i_1}, \ldots, V_k \subseteq B_{i_k}$ are open and finite intersections of open sets are open. Therefore g Ⓒ by Lemma (2.1.14).

  Let $\mathcal{A}_1, \mathcal{A}_2$ be any two topologies on A having this property. Choose $g = \mathrm{id}_A$. Now for $C = A$ and A both with topology $\mathcal{A}_1$ we find that as $\mathrm{id}_A : A \to A$ Ⓒ, so must all $f_i \circ \mathrm{id}_A = f_i$ be Ⓒ for A with topology $\mathcal{A}_1$, and similarly with topology $\mathcal{A}_2$. Consider $C = A$ with topology $\mathcal{A}_1$ and A with topology $\mathcal{A}_2$, then as all $f_i \circ \mathrm{id}_A = f_i$ Ⓒ, id_A Ⓒ. But this implies that $\mathcal{A}_1 \supseteq \mathcal{A}_2$. Do the same with both topologies interchanged to obtain that $\mathcal{A}_1 = \mathcal{A}_2$: the topology on A is uniquely determined by this property.

$\square$

Definition 2.1.20: Subspace topology
Let A $\boxed{\mathrm{T}}$, and $B \subseteq A$ any subset.

Then we will, unless specified otherwise, consider B $\boxed{\mathrm{T}}$ having the initial topology of the inclusion map $f : B \to A : a \mapsto a$.

From Definition (2.1.18) we see that the topology on $B \subseteq A$ consists precisely of all sets $B \cap U$ where $U \subseteq A$ is open.

Definition 2.1.21: Product topology
Let $\{A_i | i \in I\}$ be a collection of topological spaces.

Then we will, unless specified otherwise, consider the product $\prod_{i \in I} A_i$ $\boxed{\mathrm{T}}$ having the initial topology of the projection maps $\{f_i : \prod_{j \in I} A_j \to A_i : g \mapsto g(i) | i \in I\}$.

Example 2.1.22: $\mathbb{R}^k$
Let $k \in \mathbb{N}$. Then the product topology on $\mathbb{R}^k = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{k} = \prod_{i \in \{1, \ldots, k\}} \mathbb{R}$ is generated (Lemma (2.1.19)) by sets

$$]x_1, y_1[\times \ldots \times]x_k, y_k[$$

where all $x_i, x_i \in \mathbb{R}$, $x_i < y_i$ for $1 \leq i \leq k$, since the collection $\{]x, y[\subseteq \mathbb{R} | x, y \in \mathbb{R}, x < y\}$ forms a topological basis which generates the topology on $\mathbb{R}$.

Definition 2.1.23: Final topology

Let A be a set, $\{f_i : B_i \to A | i \in I\}$ a collection of maps with B_i [T] for all $i \in I$.

Then we define the *final topology on A with respect to* $\{f_i : B_i \to A | i \in I\}$ as

$$\{U \subseteq A \mid \forall i \in I : f_i^{-1}(U) \text{ open in } B_i\}.$$

This is indeed a topology: $f_i^{-1}(A) = B_i \subseteq B_i$, $f_i^{-1}(\emptyset) = \emptyset \subseteq B_i$ are both open, so A and $\emptyset$ are part of the final topology. If all of $\{U_j \subseteq A | j \in J\}$ are part of the final topology, then $f_i^{-1}(\bigcup_{j \in J} U_j) = \bigcup_{j \in J} f_i^{-1}(U_j) \subseteq B_i$ is open, because all $f_i^{-1}(U_j) \subseteq B_i$ are open by assumption. Hence $\bigcup_{j \in J} U_j$ is part of the final topology. If U_1, U_2 are part of the final topology, then $f_i^{-1}(U_1 \cap U_2) = f_i^{-1}(U_1) \cap f_i^{-1}(U_2) \subseteq B_i$ is open because $f_i^{-1}(U_1), f_i^{-1}(U_2) \subseteq B_i$ are open by assumption. Hence $U_1 \cap U_2$ is part of the final topology.

So the final topology is a topology.

Lemma 2.1.24: Properties of the final topology

Let A be a set, $\{f_i : B_i \to A | i \in I\}$ a collection of maps with B_i [T] for all $i \in I$.

Let the topology on A be the final topology with respect to this collection.

- The final topology is the unique largest topology for which f_i (C) for all $i \in I$.
- For any C [T] and $g : A \to C$ we have that g (C) if and only if for all $i \in I$, $g \circ f_i : B_i \to C$ (C). The final topology is the unique topology on A with this property.

Proof. For the first item, let $\mathcal{A}$ be any topology on A for which all f_i (C). Then for any $U \in \mathcal{A}$, $f_i^{-1}(U) \subseteq B_i$ is open, but then U is an element from the final topology. Therefore $\mathcal{A}$ is contained in the final topology. So the final topology is the largest topology for which all f_i (C) and by being the largest, it is also unique.

Looking at Lemma (2.1.14), it is clear that all f_i (C) with respect to the final topology.

For the second item, let C [T] and $g : A \to C$ be given. Suppose g (C), then as all f_i (C) we have (Lemma (2.1.10)) that all $g \circ f_i$ (C).

Suppose g not (C), then by Lemma (2.1.14) there exists a $W \subseteq C$ open such that $g^{-1}(W) \subseteq A$ is not open. By definition of the final topology, we therefore obtain an $i \in I$ such that $f_i^{-1}(g^{-1}(W)) \subseteq B_i$ is not open, but then $(g \circ f_i)^{-1}(W) = f_i^{-1}(g^{-1}(W)) \subseteq B_i$ is not open, while $W \subseteq C$ is open. So for this i, $g \circ f_i$ not (C). Now take the contrapositive to obtain that if for all $i \in I$, $g \circ f_i$ (C), then g (C).

Uniqueness is proven in the same way as for the initial topology. $\square$

Definition 2.1.25: (Disjoint) union topology

Let $\{A_i | i \in I\}$ be a collection of topological spaces.

Then we will, unless specified otherwise, consider the disjoint union $\coprod_{i \in I} A_i$ [T] having the initial topology of the collection $\{f_i : A_i \to \coprod_{j \in I} A_j : a \mapsto (i, a) | i \in I\}$.

We will consider the union $\bigcup_{i \in I} A_i$ [T] having the initial topology of the collection $\{f_i : A_i \to \bigcup_{j \in I} A_j : a \mapsto a\}$.

If we can write $A = \bigcup_{i\in I} A_i$, then we can consider for each $i \in I$, $A_i \subseteq A$ with the subspace topology (Definition (2.1.20)). Then the union topology (Definition (2.1.25)) of $\bigcup_{i\in I} A_i$ is the same as the topology of A if all the $A_i \subseteq A$ are open. This is not necessarily true in general: consider for example $\bigcup_{a\in A}\{a\}$ with the union topology induced by all subspaces $\{a\} \subseteq A$, then any subset of $\bigcup_{a\in A}\{a\}$ is open.

Definition 2.1.26: Quotient topology
Let A $\boxed{\mathrm{T}}$, B a set and $f : A \to B$ a surjective function.

Then the *quotient topology on B with respect to f* is the final topology of what is called the *quotient map* $f : A \to B$.

Note that any surjective function $f : A \to B$ gives rise to an equivalence relation $\sim$ on A by letting $a_1 \sim a_2$ if and only if $f(a_1) = f(a_2)$ for all $a_1, a_2 \in A$. Conversely any equivalence relation $\sim$ on A gives rise to a surjective function $f : A \to A/\sim: a \mapsto \{a_1 \in A | a \sim a_1\}$. It is easily verified that both these formulations are equivalent.

Example 2.1.27: Möbius strip
Let $A = [0, 1] \times [0, 1]$ with the subspace topology from $\mathbb{R}^2$. Then we can define an equivalence relation by letting $(0, y) \sim (1, 1-y)$ for all $y \in [0, 1]$ and $(x, y) \sim (x, y)$ for all $(x, y) \in [0, 1] \times [0, 1]$. The quotient space $A/\sim$ together with the quotient topology is what is known as the *Möbius strip* (we glue the $x = 0$ and $x = 1$ ends of A together after twisting them for 180 degrees by relating $(0, y)$ to $(1, 1 - y)$).

The initial and final topology can neatly be expressed as the unique topologies needed for continuity of all maps in the commutative diagrams:

Initial

$$\begin{tikzcd} C \arrow[r, "g"] \arrow[dr, "f_i\circ g"'] & A \arrow[d, "f_i"] \\ & B_i \end{tikzcd}$$

Final

$$\begin{tikzcd} B_i \arrow[r, "f_i"] \arrow[dr, "g\circ f_i"'] & A \arrow[d, "g"] \\ & C \end{tikzcd}$$

for all $i \in I$.

Theorem 2.1.28: Operations preserving continuity
Let A, B $\boxed{\mathrm{T}}$.

Composition: let C $\boxed{\mathrm{T}}$, $a \in A$, $f : A \to B$, $g : B \to C$.
If f C a and g C $f(a)$, then $g \circ f$ C a.
In particular if f C and g C, then $g \circ f$ C.

Glueing: let $U_1, U_2 \subseteq A$ be both open or both closed such that $A = U_1 \cup U_2$ and let $f_1 : U_1 \to B$, $f_2 : U_2 \to B$.
If f_1, f_2 C with respect to the subspace topologies on U_1 and U_2, and $f_1(a) = f_2(a)$ for all $a \in U_1 \cap U_2$, then there exists a unique function $f : A \to B$ C such that $f_1 = f|_{U_1}$ and $f_2 = f|_{U_2}$.

Restricting domain: let $C \subseteq A$ a subset, and $f : A \to B$.
If f C, then $f|_C$ C where C has the subspace topology.

Expanding image: suppose $B \subseteq C$ with the subspace topology and let $f : A \to B$. If f Ⓒ, then $f : A \to C$ Ⓒ.

Restricting image: let $f : A \to B$, and $C \subseteq B$ a subset with $f(A) \subseteq C$. If f Ⓒ, then $f : A \to C$ Ⓒ where C has the subspace topology.

Constantness: let $f : A \to B$ and suppose there exists a $b \in B$ such that $f(a) = b$ for all $a \in A$, then f Ⓒ.

Proof. We will frequently use Lemma (2.1.14) in this proof.

- Follows directly from Lemma (2.1.10).
- Let $f_1 : U_1 \to B$, $f_2 : U_2 \to B$ both Ⓒ be given for $A = U_1 \cup U_2$ and suppose $f_1(a) = f_2(a)$ for all $a \in U_1 \cap U_2$. Define $f : A \to B$ by $f(a) := f_1(a)$ if $a \in U_1$ and $f(a) := f_2(a)$ otherwise. Then $f|_{U_1} = f_1$, $f|_{U_2} = f_2$ by definition. Note that for any $V \subseteq B$ we have $f^{-1}(V) = f_1^{-1}(V) \cup f_2^{-1}(V)$.

  Suppose U_1 and U_2 are open, let $V \subseteq B$ be open. Then as f_1, f_2 Ⓒ, we have $f_1^{-1}(V) \subseteq U_1$, $f_2^{-1}(V) \subseteq U_2$ are open. Therefore (subspace topology) $f_1^{-1}(V) = U_3 \cap U_1$, $f_2^{-1}(V) = U_4 \cap U_2$ for some $U_3, U_4 \subseteq A$ open. Hence $f^{-1}(V) = f_1^{-1}(V) \cup f_2^{-1}(V) = (U_3 \cap U_1) \cup (U_4 \cap U_2)$ which is open as U_1 and U_2 are open. If U_1 and U_2 are closed, we can follow the same route for $V \subseteq B$ closed to obtain that $f^{-1}(V)$ is closed. Therefore f Ⓒ. Uniqueness follows directly from the demand that $f|_{U_1} = f_1$, $f|_{U_2} = f_2$ together with $A = U_1 \cup U_2$.
- Let $C \subseteq A$ and suppose $f : A \to B$ Ⓒ. Let $V \subseteq B$ be open, then $(f|_C)^{-1}(V) = f^{-1}(V) \cap C$ which is open in the subspace topology, because $f^{-1}(V)$ is open. Therefore $f|_C$ Ⓒ.
- Suppose $B \subseteq C$ and $f : A \to B$ Ⓒ. Then $f : A \to C$ is the composition of f with the inclusion map $B \to C$ which is Ⓒ by choice of the subspace topology, therefore $f : A \to C$ Ⓒ by the first item.
- Let $f : A \to B$ Ⓒ, $f(A) \subseteq C \subseteq B$. Let $W \subseteq C$ be any open set. Because of the subspace topology $W = C \cap V$ for $V \subseteq B$ open. As $f(A) \subseteq C$ we have $f^{-1}(W) = f^{-1}(C \cap V) = A \cap f^{-1}(V) = f^{-1}(V)$, which is open because $f : A \to B$ Ⓒ. Therefore $f : A \to C$ Ⓒ.
- Let $f : A \to B$ satisfy $f(a) = b$ for all $a \in A$. Then for any open $V \subseteq B$ we have that $f^{-1}(V)$ equals either $\emptyset$ ($b \notin V$) or A ($b \in V$) which are both part of the topology of A by definition. So f Ⓒ.

□

Definition 2.1.29: Morphisms of topological spaces

Let A, B $\boxed{\mathrm{T}}$.

Then all maps $f : A \to B$ Ⓒ are *morphisms between A and B*.

The *identity morphism* is the continuous map

$$\mathrm{id}_A : A \to A : a \mapsto a.$$

Topological isomorphisms (denoted by $\boxed{\mathrm{T}}$-isomorphisms) are commonly called *homeomorphisms*.

2.2 Separation axioms

This definition deals with the increasing precision with which we can distinguish subsets of a topological space using the space's topology.

Definition 2.2.1: Separation ($\boxed{\text{T0}}$, $\boxed{\text{T1}}$, $\boxed{\text{T2}}$, $\boxed{\text{T3}}$, $\boxed{\text{T3.5}}$, $\boxed{\text{T4}}$, $\boxed{\text{T6}}$)
Let A $\boxed{\text{T}}$.

The following properties should hold for any $a_1, a_2 \in A$, $a_1 \neq a_2$ (separation of distinct points in A).

- We say A $\boxed{\text{T0}}$ if
$$\exists U \subseteq A \text{ open } : a_1 \in U \wedge a_2 \notin U.$$
- We say A $\boxed{\text{T1}}$ if
$$\exists U_1, U_2 \subseteq A \text{ open } : a_1 \in U_1 \wedge a_2 \in U_2 \wedge a_1 \notin U_2 \wedge a_2 \notin U_1.$$
- We say A $\boxed{\text{T2}}$ or *Hausdorff* if
$$\exists U_1, U_2 \subseteq A \text{ open } : a_1 \in U_1 \wedge a_2 \in U_2 \wedge U_1 \cap U_2 = \emptyset.$$

The following properties should hold for any $a \in A$, $B \subseteq A$ closed, $a \notin B$ (separation of a point and a closed set).

- We say A *is regular* if
$$\exists U_1, U_2 \subseteq A \text{ open } : a \in U_1 \wedge B \subseteq U_2 \wedge U_1 \cap U_2 = \emptyset.$$
- We say A $\boxed{\text{T3}}$ if A is regular and $\boxed{\text{T2}}$.
- We say A *is completely regular* if
$$\exists f : A \to \mathbb{R} \boxed{\text{C}} : f(a) = 0 \wedge f(B) = \{1\}.$$
- We say A $\boxed{\text{T3.5}}$ or *Tychonoff* if A is completely regular and $\boxed{\text{T2}}$.

The following properties should hold for any $B, C \subseteq A$ closed, $B \cap C = \emptyset$ (separation of distinct closed sets).

- We say A *is normal* if
$$\exists U_1, U_2 \subseteq A \text{ open } : B \subseteq U_1 \wedge C \subseteq U_2 \wedge U_1 \cap U_2 = \emptyset.$$
- We say A $\boxed{\text{T4}}$ if A is normal and $\boxed{\text{T2}}$.
- We say A *is perfectly normal* if
$$\exists f : A \to \mathbb{R} \boxed{\text{C}} : f^{-1}(\{0\}) = B \wedge f^{-1}(\{1\}) = C.$$
- We say A $\boxed{\text{T6}}$ if A is perfectly normal and $\boxed{\text{T2}}$.

Lemma 2.2.2: Closed point sets
Let A $\boxed{\text{T}}$.

Then A $\boxed{\text{T1}}$ if and only if for all $a \in A$, $\{a\} \subseteq A$ is closed.

Proof. Suppose A T1. Let $a_1 \in A \setminus \{a\}$, then $a_1 \neq a$ so (A is T1) there exist $U, U_{a_1} \subseteq A$ open such that $a \in U$, $a_1 \in U_{a_1}$, $a \notin U_{a_1}$, $a_1 \notin U$. So for all $a_1 \in A \setminus \{a\}$ there exists a $U_{a_1} \subseteq A$ open such that $a_1 \in U_{a_1} \subseteq A \setminus \{a\}$. But then $A \setminus \{a\} = \bigcup_{a_1 \in A \setminus \{a\}} \{a_1\} \subseteq \bigcup_{a_1 \in A \setminus \{a\}} U_{a_1} \subseteq A$, so $A = \bigcup_{a_1 \in A \setminus \{a\}} U_{a_1}$ which is open as all U_{a_1} are open. Therefore $\{a\} = A \setminus (A \setminus \{a\})$ is closed.

Suppose $\{a\} \subseteq A$ is closed for all $a \in A$. Let $a_1, a_2 \in A$, $a_1 \neq a_2$. By assumption $\{a_1\}, \{a_2\} \subseteq A$ are closed, so $U_1 := A \setminus \{a_2\}$, $U_2 := A \setminus \{a_1\}$ are open and $a_1 \in U_1$, $a_2 \in U_2$, $a_1 \notin U_2$, $a_2 \notin U_1$, because $a_1 \neq a_2$. So A T1. □

Lemma 2.2.3: Shrinking

Let A T.

Then A is normal if and only if for all $B \subseteq U_1 \subseteq A$ where B is closed and U_1 is open there exists an open $U_2 \subseteq A$ such that $B \subseteq U_2 \subseteq \overline{U}_2 \subseteq U_1$.

Proof. Suppose A is normal and let $B \subseteq U_1 \subseteq A$ be given. As U_1 is open, $C := A \setminus U_1 \subseteq A$ is closed, so because A is normal and $B, C \subseteq A$ are closed and disjoint there exist open sets $U_2, U_3 \subseteq A$ such that $B \subseteq U_2$ and $C \subseteq U_3$ and $U_2 \cap U_3 = \emptyset$. Now $A \setminus U_3$ is a closed set containing U_2, so $\overline{U}_2 \subseteq A \setminus U_3 \subseteq A \setminus C = U_1$, therefore $B \subseteq U_2 \subseteq \overline{U}_2 \subseteq U_1$.

Suppose the converse holds. Let $B, C \subseteq A$ be arbitrary closed sets for which $B \cap C = \emptyset$. Then $U_1 := A \setminus C$ is an open set containing B, so by assumption there exists an open $U_2 \subseteq A$ such that $B \subseteq U_2 \subseteq \overline{U}_2 \subseteq U_1$, but then U_2 and $A \setminus \overline{U}_2$ are disjoint open sets containing B and C respectively. Since this is true for all such B and C, A is normal. □

The following theorem explains why there is no analogy of T3.5 for normal spaces.

Lemma 2.2.4: Urysohn's lemma

Let A T be normal.

Then for all $B, C \subseteq A$ that are closed and disjoint there exists an $f : A \to [0,1]$ C such that $f(B) = \{0\}$ and $f(C) = \{1\}$.

Proof. We follow [Mun2000]. As $\mathbb{Q} \cap [0,1] \subseteq \mathbb{Q}$ is countable, there exists a bijection $q : \mathbb{N} \to \mathbb{Q} \cap [0,1] : k \mapsto q_k$ for which $q_0 = 0$ and $q_1 = 1$. Now construct $U_{q_k} \subseteq A$ open by induction on k, such that for all $r, s \in \mathbb{Q} \cap [0,1]$ we have $r < s \to \overline{U}_r \subseteq U_s$.

First $k = 0, 1$. Define $U_1 := A \setminus C$ which is open and contains B. By Lemma (2.2.3) there exists an open set, which we will define to be U_0, such that $B \subseteq U_0 \subseteq \overline{U}_0 \subseteq U_1$. Therefore the U_{q_k} satisfy the induction hypothesis for $0 \leq k \leq 1$.

Now suppose we have constructed the U_{q_k} with the inclusion property for $0 \leq k \leq k_0$. As $\{q_0, \ldots, q_{k_0}\}$ is finite, there exist $l, m \in \{0, \ldots, k_0\}$ such that $q_l < q_{k_0+1} < q_m$ with $|q_l - q_{k_0+1}|$ and $|q_{k_0+1} - q_m|$ minimal. By the induction hypothesis $\overline{U}_{q_l} \subseteq U_{q_m}$, so using Lemma (2.2.3) we obtain $U_{q_{k_0+1}} \subseteq A$ as the open set for which $\overline{U}_{q_l} \subseteq U_{q_{k_0+1}} \subseteq \overline{U}_{q_{k_0+1}} \subseteq U_{q_m}$. Through choice of l and m, the U_{q_k} satisfy the induction hypothesis for $0 \leq k \leq k_0 + 1$. By induction this permits us to construct the U_{q_k} for all $k \in \mathbb{N}$ with the desired inclusion property, and therefore U_r for all $r \in \mathbb{Q} \cap [0,1]$ as $q : \mathbb{N} \to \mathbb{Q} \cap [0,1]$ is a bijection.

Define $U_r := \emptyset$ for $r < 0$ and $U_r := A$ for $r > 1$ to obtain open $U_r \subseteq A$ for all $r \in \mathbb{Q}$ satisfying $\overline{U}_r \subseteq U_s$ whenever $r < s$. Now define $f : A \to [0,1]$ by

$$f(a) := \inf\{r \in \mathbb{Q} | a \in U_r\}$$

which indeed lies within $[0,1]$ by choice of the U_r. As $U_0 \supseteq B$ and $U_r = \emptyset$ for $r < 0$ we see that $f(B) = \{0\}$, because $U_1 = A \setminus C$ and $U_r = A$ for all $r > 1$ we have $f(C) = \{1\}$.

Because of the inclusion property, if $a \in \overline{U}_r$, then $\{s \in \mathbb{Q} | a \in U_s\} \subseteq [r, \infty[$, so $f(a) \leq r$, and if $a \notin U_r$, then $a \notin U_s$ for all $s < r$ and hence $f(a) \geq r$.

Let $a \in A$ be fixed and $\epsilon \in]0, \infty[$. Choose $r, s \in \mathbb{Q}$ such that $f(a) - \epsilon < r < f(a) < s < f(a) + \epsilon$ (such r and s exist because $\mathbb{Q}$ is dense in $\mathbb{R}$) and let $U := U_s \setminus \overline{U}_r \subseteq A$ open, then $r < f(a) < s$ implies $a \notin \overline{U}_r$ and $a \in U_s$, so $a \in U$. Furthermore, for all $a_1 \in U$ we have $a_1 \in \overline{U}_s$ and $a_1 \notin U_r$, so $f(a_1) \in [r, s] \subseteq]f(a) - \epsilon, f(a) + \epsilon[$. Because this is true for all $\epsilon \in]0, \infty[$, f c a. As $a \in A$ was arbitrary, f c. □

Theorem 2.2.5: Relations between the separation axioms

Let A $\boxed{\text{T}}$.

Then

- perfectly normal $\Rightarrow$ normal,
- $\boxed{\text{T1}}$ and normal $\Rightarrow$ completely regular,
- completely regular $\Rightarrow$ regular,
- $\boxed{\text{T6}} \Rightarrow \boxed{\text{T4}} \Rightarrow \boxed{\text{T3.5}} \Rightarrow \boxed{\text{T3}} \Rightarrow \boxed{\text{T2}} \Rightarrow \boxed{\text{T1}} \Rightarrow \boxed{\text{T0}}$.

Proof.

- Suppose A is perfectly normal and let $B, C \subseteq D$ be closed and disjoint. As A is perfectly normal there exists a $f : A \to \mathbb{R}$ c such that $f^{-1}(\{0\}) = B$, $f^{-1}(\{1\}) = C$. Since $]-\infty, 1/2[,]1/2, \infty[\subseteq \mathbb{R}$ are open and disjoint, so are $f^{-1}(]-\infty, 1/2[), f^{-1}(]1/2, \infty[) \subseteq A$ because f c. As $0 \in]-\infty, 1/2[$, $1 \in]1/2, \infty[$ we see that $B \subseteq f^{-1}(]-\infty, 1/2[)$, $C \subseteq f^{-1}(]1/2, \infty[)$. So we can separate closed, disjoint B and C with open sets: A is normal.

- Let A be normal and $\boxed{\text{T1}}$. Let $a \in A$ and $B \subseteq A$ closed such that $a \notin B$. Because A $\boxed{\text{T1}}$, by Lemma (2.2.2) $\{a\} \subseteq A$ is closed and disjoint from B, therefore (A is normal by assumption) by Lemma (2.2.4) there exists an $f : A \to \mathbb{R}$ c with $f(\{a\}) = \{0\}$ and $f(B) = \{1\}$. Therefore A is completely regular.

- Suppose A is completely regular. Use the trick from the reduction from perfectly normal to normal to obtain regularity of A.

- It is clear from Definition (2.2.1) that $\boxed{\text{T2}} \Rightarrow \boxed{\text{T1}} \Rightarrow \boxed{\text{T0}}$, so for the final item, simply combine all of the above, noting that for $\boxed{\text{T3}}$ and above we assume $\boxed{\text{T2}}$ (and hence have $\boxed{\text{T1}}$) by definition.

□

Lemma 2.2.6: Uniqueness of limits

Let A $\boxed{\text{T}}$, B $\boxed{\text{T}}\boxed{\text{T2}}$, $f : A \to B$, $a \in A$, and $b_1, b_2 \in B$.

If $\lim_{x \to a} f(x) = b_1$ and $\lim_{x \to a} f(x) = b_2$, then $b_1 = b_2$.

Proof. Suppose $b_1 \neq b_2$, then because B $\boxed{\text{T2}}$ there exist open neighbourhoods V_1 and V_2 of b_1 and b_2 respectively in B such that $V_1 \cap V_2 = \emptyset$. Because $\lim_{x \to a} f(x) = b_1$ there exists an open neighbourhood U_1 of a in A such that $f(U_1) \subseteq V_1$ and because $\lim_{x \to a} f(x) = b_2$ there exists an open neighbourhood

U_2 of a in A such that $f(U_2) \subseteq V_2$. Now $f(U_1 \cap U_2) \subseteq V_1 \cap V_2 = \emptyset$, however $a \in U_1 \cap U_2$ so $f(a) \in V_1 \cap V_2$, leading to a contradiction.

Therefore necessarily $b_1 = b_2$. $\square$

This permits us in a Hausdorff space to actually talk about *the* limit of a certain function at a certain point.

Lemma 2.2.7: Graphs of C functions are closed

Let A T, B T T2, $f : A \to B$.

If f C, then

$$\operatorname{graph}(f) := \{(a,b) \in A \times B \mid b = f(a)\} \subseteq A \times B$$

is closed.

Proof. Let $(a,b) \in A \times B \setminus \operatorname{graph}(f)$, then $b \neq f(a)$, so as B T2 there exist open neighbourhoods V_1, V_2 of b resp. $f(a)$ in B such that $V_1 \cap V_2 = \emptyset$. As f C, there exists an open neighbourhood U_2 of a in A with $f(U_2) \subseteq V_2$. Now $V_{(a,b)} := U_2 \times V_1$ is an open neighbourhood of $(a,b) \in A \times B$ and for any $(a_1, b_1) \in V_{(a,b)}$ we have $f(a_1) \in f(U_2) \subseteq V_2$, so as $b_1 \in V_1$ and $V_1 \cap V_2 = \emptyset$, we find $f(a_1) \neq b_1$ and hence $(a_1, b_1) \notin \operatorname{graph}(f)$.

So for any $(a,b) \in A \times B \setminus \operatorname{graph}(f)$ there exists an open neighbourhood $V_{(a,b)}$ of (a,b) in $A \times B$ such that $V_{(a,b)} \cap \operatorname{graph}(f) = \emptyset$. Therefore $A \times B \setminus \operatorname{graph}(f) = \bigcup_{(a,b) \in A \times B \setminus \operatorname{graph}(f)} \{(a,b)\} \subseteq \bigcup_{(a,b) \in A \times B \setminus \operatorname{graph}(f)} V_{(a,b)} \subseteq A \times B \setminus \operatorname{graph}(f)$ which is open as a union of a collection of open subsets. Hence $\operatorname{graph}(f) \subseteq A \times B$ is closed. $\square$

A partial converse to this result is given in Lemma (2.2.8) and Theorem (4.4.4).

Lemma 2.2.8

Let A, B T, $f : A \to B$.

If $A \times B \to A : (a,b) \mapsto a$ is a closed map and $\operatorname{graph}(f) \subseteq A \times B$ is closed, then f C.

Proof. Define $g_1 : A \times B \to A : (a,b) \mapsto a$, $g_2 : A \times B \to B : (a,b) \mapsto b$. Then g_1, g_2 C by definition of the initial topology on $A \times B$. Now for any $V \subseteq B$ closed we have $f^{-1}(V) = \{a \in A | f(a) \in V\} = \{g_1(a, f(a)) \in A | g_2(a, f(a)) \in V\} = \{g_1(a,b) \in A | (a,b) \in \operatorname{graph}(f), g_2(a,b) \in V\} = g_1(\operatorname{graph}(f) \cap g_2^{-1}(V))$. As g_2 C (Lemma (2.1.14)), $g_2^{-1}(V)$ is closed, $\operatorname{graph}(f)$ is closed by assumption, and g_1 is a closed map by assumption, we have that therefore $f^{-1}(V) \subseteq A$ is closed. Since this is true for all $V \subseteq B$ closed, f C by Lemma (2.1.14). $\square$

2.3 Sequences

Definition 2.3.1: Sequence

Let A T.

Then a *sequence* is a map $x : \mathbb{N} \to A : k \mapsto x_k$. We say that x *has limit a in A*, denoted by

$$\lim_{k \to \infty} x_k = a$$

if for all open neighbourhoods U of a in A there exists a $k \in \mathbb{N}$ such that for all $l \geq k$ we have $x_l \in U$.

If there exists an $a \in A$ such that $\lim_{k\to\infty} x_k = a$ we say that *the sequence x is convergent in A.*

A sequence $x' : \mathbb{N} \to A$ is called a *subsequence of $x : \mathbb{N} \to A$* if there exist $k_1 < k_2 < \ldots$ in $\mathbb{N}$ such that $x'_l = x_{k_l}$ for all $l \in \mathbb{N}$.

Note that if a sequence is convergent, then so is any subsequence and for any fixed $k \in \mathbb{N}$, the sequence $l \mapsto x_l$ is convergent if and only if $l \mapsto x_{k+l}$ is convergent.

Example 2.3.2: $\hat{\mathbb{N}}$

$\hat{\mathbb{N}}$ is the topological space defined as a set as

$$\hat{\mathbb{N}} := \mathbb{N} \cup \{\infty\},$$

with topology consisting of all subsets

$$\{k, k+1, k+2, \ldots\} \cup \{\infty\}$$

of $\hat{\mathbb{N}}$ for all $k \in \mathbb{N}$, together with the empty set. That this is a topology follows from the fact that $\mathbb{N}$ is a well-ordered set: each non-empty subset of $\mathbb{N}$ has a least element.

Let A $\boxed{\mathrm{T}}$, then looking at Definition (2.3.1) we see that any map

$$x : \hat{\mathbb{N}} \to A : k \mapsto x(k) = x_k$$

satisfies $x \; \boxed{\mathrm{C}} \; \infty$ if and only if $\lim_{k\to\infty} x(k) = x(\infty)$ if and only if the *sequence* $x|_{\mathbb{N}}$ satisfies $\lim_{k\to\infty}(x|_{\mathbb{N}})_k = x(\infty)$ for the point $x(\infty) \in A$.

Definition 2.3.3: First countability ($\boxed{\mathrm{1C}}$)

Let A $\boxed{\mathrm{T}}$.

Then we say that A is *first countable* (denoted by A $\boxed{\mathrm{1C}}$) if for all $a \in A$ there exists a basis of open neighbourhoods of a in A that is countable (cf. Definition (2.1.4)).

Note that we may, without loss of generality, suppose this countable collection $U_1, U_2, \ldots$ of open neighbourhoods to be descending: $U_1 \supseteq U_2 \supseteq \ldots$ by considering U_1, $U_1 \cap U_2$, $U_1 \cap U_2 \cap U_3$, $\ldots$, which are all open neighbourhoods of a in A.

Lemma 2.3.4

Let A $\boxed{\mathrm{T}}\boxed{\mathrm{1C}}$, B $\boxed{\mathrm{T}}$, and $f : A \to B$.

Let $a \in A$ and $b \in B$, then

$$\lim_{x\to a} f(x) = b$$

if and only if for all sequences $x : \mathbb{N} \to A$ satisfying $\lim_{k\to\infty} x_k = a$, we have

$$\lim_{k\to\infty} f(x_k) = b.$$

Proof. Suppose $\lim_{x \to a} f(x) = b$ and let $x : \mathbb{N} \to A$ be any sequence satisfying $\lim_{k\to\infty} x_k = a$. Then $\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty}(f \circ x)(k) = f(b)$, because of the chain rule for limits and Example (2.3.2).

Suppose $\lim_{x\to a} f(x) \neq b$. Then there exists an open neighbourhood V of b in B such that for all open neighbourhoods U of a in A there exists a point $a_1 \in U$ for which $f(a_1) \notin V$. Because A $\boxed{\mathrm{1C}}$, there exists a descending countable collection $U_1, U_2, \ldots$ of open neighbourhoods of a in A such that for each open neighbourhood U of a in A there exists a $k \in \mathbb{N}$ such that $U_l \subseteq U$ for all $l \geq k$. Construct a sequence $x : \mathbb{N} \to A$ by mapping $k \in \mathbb{N}$ to a point $x_k \in U_k$ for which $f(x_k) \notin V$. Then $\lim_{k\to\infty} x_k = a$ because for any open neighbourhood U of a in A there exists a $k \in \mathbb{N}$ such that for all $l \geq k$ we have $U_l \subseteq U$ and since $x_l \in U_l$ for all $l \in \mathbb{N}$ this means that $\lim_{k\to\infty} x_k = a$. However, by construction, for all $k \in \mathbb{N}$, we have $f(x_k) \notin V$ and hence $\lim_{k\to\infty} f(x_k) \neq b$. So there exists a sequence $x : \mathbb{N} \to A$ with $\lim_{k\to\infty} f(x_k) \neq b$. □

Lemma 2.3.5
Let A $\boxed{\mathrm{T}}\boxed{\mathrm{1C}}$, and $B \subseteq A$ a subset.

Then $a \in \overline{B}$ if and only if there exists a convergent sequence $x : \mathbb{N} \to B$ such that $a = \lim_{k\to\infty} x_k$.

Proof. Let $a \in \overline{B}$. As A $\boxed{\mathrm{1C}}$, there exists a descending countable basis of neighbourhoods $U_1, U_2, \ldots$ of a in A. By Lemma (2.1.3), for each $k \in \mathbb{N}$, $U_k \cap B \neq \emptyset$ because U_k is an open neighbourhood of $a \in \overline{B}$. Hence we can for each $k \in \mathbb{N}$ pick an $x_k \in U_k \cap B$ to obtain a sequence $x : \mathbb{N} \to B : k \mapsto x_k$. Let U be any open neighbourhood of a in A, then there exists a $k \in \mathbb{N}$ such that $a \in U_l \subseteq U$ for all $l \geq k$ (as $U_1 \supseteq U_2 \supseteq \ldots$). Hence for all $l \geq k$, $x_l \in U$, so $\lim_{k\to\infty} x_k = a$.

Suppose conversely that $x : \mathbb{N} \to B$ with limit $a = \lim_{k\to\infty} x_k \in A$. Let U be an arbitrary open neighbourhood of a in A, then there exists a $k \in \mathbb{N}$ such that for all $l \geq k$ we have $x_l \in U$. In particular $x_k \in U \cap B$ (as $x : \mathbb{N} \to B$), so $U \cap B \neq \emptyset$. Since this is true for all open neighbourhoods U of a in A, by Lemma (2.1.3), $a \in \overline{B}$. □

2.4 Compactness

Definition 2.4.1: Collections of subsets
Let A $\boxed{\mathrm{T}}$.

A subset $\mathcal{A} \subseteq \mathcal{P}(A)$ is called a *collection of subsets of* A. A collection $\mathcal{A}$ is called

- *open* (resp. *closed*) if $U \subseteq A$ is open (resp. *closed*) for all $U \in \mathcal{A}$,
- *a cover of* A if $A = \bigcup \mathcal{A}$,
- *finite* if $\mathcal{A}$ consists of a finite number of elements,
- *locally finite* if for all $a \in A$ there exists an open neighbourhood U of a in A such that $\{U_1 \in \mathcal{A} | U_1 \cap U \neq \emptyset\}$ is finite,
- *countably locally finite* if $\mathcal{A} = \bigcup_{k\in\mathbb{N}} \mathcal{A}_k$ where each $\mathcal{A}_k$ is a locally finite collection.

For any two collections $\mathcal{A}_1$, $\mathcal{A}_2$ of subsets of A we furthermore say that $\mathcal{A}_1$ is a *subcollection* of $\mathcal{A}_2$ if $\mathcal{A}_1 \subseteq \mathcal{A}_2$, and that $\mathcal{A}_1$ is *a refinement of* $\mathcal{A}_2$ if for all $U_1 \in \mathcal{A}_1$ there exists a $U_2 \in \mathcal{A}_2$ such that $U_1 \subseteq U_2$.

Definition 2.4.2: Compactness ([Cpt])
Let A [T].

Then A is *compact* (denoted by A [Cpt]) if all open covers of A contain a finite subcollection covering A.

Lemma 2.4.3
Let A [T][Cpt], B [T].

If $f : A \to B$ [C], then $f(A) \subseteq B$ [Cpt].

Proof. Let $\mathcal{B}$ be an open cover of $f(A)$. Then (subspace topology) we can write $\mathcal{B} = \{V_i \cap f(A) | i \in I\}$ where all $V_i \subseteq B$ are open. Because f [C], for all $i \in I$, $f^{-1}(V_i \cap f(A)) = f^{-1}(V_i) \cap A \subseteq A$ is open and since $A = f^{-1}(f(A)) = f^{-1}(\bigcup \mathcal{B}) = \bigcup_{i \in I} f^{-1}(V_i \cap f(A))$ the collection $\mathcal{A} := \{f^{-1}(V_i \cap f(A)) | i \in I\}$ forms an open cover of A.

Because A is [Cpt] there exists a finite subcollection $f^{-1}(V_{i_1} \cap f(A)), \ldots, f^{-1}(V_{i_k} \cap f(A))$ of this cover, which covers A. Hence $V_{i_1} \cap f(A), \ldots, V_{i_k} \cap f(A)$ is a finite open cover of $f(A)$ which is furthermore a subcollection of $\mathcal{B}$. This makes $f(A)$ [Cpt]. □

Lemma 2.4.4
Let A [T][Cpt].

Then for any $B \subseteq A$ closed, B [Cpt].

Proof. Let $\mathcal{A}$ be any open cover of B, then because of the subspace topology we can write $\mathcal{A} = \{U_i \cap B | i \in I\}$ where all $U_i \subseteq A$ are open. Note that $\{U_i | i \in I\} \cup \{A \setminus B\}$ now forms an open cover of A, because B is closed. Now A [Cpt], so this open cover has a finite subcollection covering A which in turn also covers B. This shows that B [Cpt]. □

Lemma 2.4.5
Let A [T][T2].

Then for any $B \subseteq A$ [Cpt], B is closed.

Proof. Let $a \in A \setminus B$, then for any $b \in B$ we have $b \neq a$ and hence (A [T2]) there exists a neighbourhood U_b of a in A and V_b of b in B such that $U_b \cap V_b = \emptyset$. The collection $\{V_b | b \in B\}$ forms an open cover of B (as $b \in V_b$ for all $b \in B$), so (B [Cpt]) there exists a finite number of $b_1, \ldots, b_k \in B$ such that $B \subseteq V_{b_1} \cup \ldots \cup V_{b_k}$. Hence $a \in U_{b_1} \cap \ldots \cap U_{b_k} \subseteq A \setminus (V_{b_1} \cup \ldots \cup V_{b_k}) \subseteq A \setminus B$. So for each $a \in A \setminus B$ there exists an open neighbourhood U_a $(= U_{b_1} \cap \ldots \cap U_{b_k})$ of a in A with $a \in U_a \subseteq A \setminus B$. Therefore $A \setminus B = \bigcup_{a \in A \setminus B} \{a\} \subseteq \bigcup_{a \in A \setminus B} U_a \subseteq A \setminus B$, so $A \setminus B$ is open and hence B is closed. □

Lemma 2.4.6: Sequential compactness
Let A [T][1C][Cpt].

Then for all sequences $x : \mathbb{N} \to A$ there exists a subsequence $x' : \mathbb{N} \to A$ of x which is convergent.

Proof. Let $x : \mathbb{N} \to A$ be any sequence. Suppose there exists an $a \in A$ such that for all open neighbourhoods U of a in A the set $\{k \in \mathbb{N} | x_k \in U\}$ is infinite. Since A 1C, a admits a countable basis of open neighbourhoods $U_1 \supseteq U_2 \supseteq \ldots$. With this basis we can construct a subsequence of x by choosing $x'_l := x_{k_l}$ for k_l the least element of the infinite set $\{k \in \mathbb{N} | x_k \in U_l\} \subseteq \mathbb{N}$ (possible as $\mathbb{N}$ is well-ordered). Now let U be an arbitrary open neighbourhood of a in A, then there exists an $m \in \mathbb{N}$ such that $a \in U_m \subseteq U$. For all $l \geq m$ we have $x'_l = x_{k_l} \in U_l \subseteq U_m \subseteq U$ by construction. Therefore $\lim_{l\to\infty} x'_l = a$ and x' is a convergent subsequence of x.

Now suppose that this is not the case: suppose that for all $a \in A$ there exists an open neighbourhood U_a of a in A such that $\{k \in \mathbb{N} | x_k \in U_a\}$ is finite, denote the number of elements of this set by $k_a \in \mathbb{N}$. The collection $\mathcal{U} := \{U_a \subseteq A | a \in A\}$ is an open cover of A, therefore (A Cpt) there exists a finite number of points $a_1, \ldots, a_l \in A$ such that $A = U_{a_1} \cup \ldots \cup U_{a_l}$. As $\{x_1, x_2, \ldots\} \subseteq A = U_{a_1} \cup \ldots \cup U_{a_l}$ and $\{x_1, x_2, \ldots\} \cap U_{a_m}$ has k_{a_m} elements for all $1 \leq m \leq l$, the set $\{x_1, x_2, \ldots\}$ has at most $k_{a_1} + \ldots + k_{a_l}$ elements and is therefore finite. Because $\mathbb{N}$ is infinite, this means that there exists some $a \in A$ such that the set $\{k \in \mathbb{N} | x_k = a\}$ is infinite. Therefore the constant sequence $x'_l := a$ for all $l \in \mathbb{N}$ is a convergent subsequence of x. □

2.5 Metric spaces

Definition 2.5.1: Metric space (d(.,.))
Let A be a set.

Then a *pseudometric on* A is a map $d : A \times A \to \mathbb{R}$, satisfying for all $a_1, a_2, a_3 \in A$ that

- $d(a_1, a_2) \geq 0$,
- $d(a_1, a_1) = 0$,
- $d(a_1, a_2) = d(a_2, a_1)$,
- $d(a_1, a_3) \leq d(a_1, a_2) + d(a_2, a_3)$.

If in addition $d(a_1, a_2) = 0 \to a_1 = a_2$, then d is called a *metric on* A.

A *(pseudo)metric space* A is a set A together with a (pseudo)metric d.

We denote the fact that A is a metric space by A d(.,.).

Definition 2.5.2: Open ball
Let A be a (pseudo)metric space.

Then for any $a \in A$ and $\delta \in]0, \infty[$ we define the *open ball of radius δ around a in A* to be

$$B_A(a, \delta) := \{a_1 \in A \mid d(a, a_1) < \delta\}.$$

Lemma 2.5.3
Let A be a (pseudo)metric space.

Then the collection of all open balls in A,

$$\mathcal{A} := \{B_A(a, \delta) \subseteq A \mid a \in A, \delta \in]0, \infty[\}$$

forms a topological basis of A.

Proof. First note that $d(a,a) = 0 < \delta$ for all $\delta \in]0,\infty[$, so $a \in B_A(a,\delta)$ for all $a \in A$ and $\delta \in]0,\infty[$. Therefore $A = \bigcup \mathcal{A}$.

Let $a_1, a_2 \in A$ and $\delta_1, \delta_2 \in]0,\infty[$. If $a_1 = a_2$ then $B_A(a_1,\delta_1) \cap B_A(a_2,\delta_2) = B_A(a_1, \min\{\delta_1,\delta_2\})$ which is again an element of the basis, so we may suppose $a_1 \neq a_2$.

Let $a_3 \in B_A(a_1,\delta_1) \cap B_A(a_2,\delta_2)$ be arbitrary. Choose $\delta_3 := \min\{\delta_1 - d(a_1,a_3), \delta_2 - d(a_1,a_3)\} > 0$, then for any $a_4 \in B_A(a_3,\delta_3)$ we have $d(a_1,a_4) \leq d(a_1,a_3) + d(a_3,a_4) < d(a_1,a_3) + \delta_3 \leq d(a_1,a_3) + \delta_1 - d(a_1,a_3) = \delta_1$, so $a_4 \in B_A(a_1,\delta_1)$. Similarly $a_4 \in B_A(a_2,\delta_2)$, so $B_A(a_3,\delta_3) \subseteq B_A(a_1,\delta_1) \cap B_A(a_2,\delta_2)$ and such a basis element exists for all $a_3 \in B_A(a_1,\delta_1) \cap B_A(a_2,\delta_2)$. Therefore $\mathcal{A}$ is a topological basis. □

Definition 2.5.4: Topology of a metric space

Let A be a (pseudo)metric space.

Then we always consider A $\boxed{\mathrm{T}}$ with topology generated by the topological basis from Lemma (2.5.3).

We call a B $\boxed{\mathrm{T}}$ *(pseudo)metrisable* if B is $\boxed{\mathrm{T}}$-isomorphic to some (pseudo)metric space A.

Note that this in particular makes all $B_A(a,\delta)$ open subsets of A.

Lemma 2.5.5

Let A, B both be (pseudo)metric spaces, $a \in A$, $b \in B$, $f : A \to B$.

Then $\lim_{x\to a} f(x) = b$ if and only if

$$\forall \epsilon \in]0,\infty[: \exists \delta \in]0,\infty[: \forall a_1 \in A : (d_A(a,a_1) < \delta \to d_B(b, f(a_1)) < \epsilon).$$

Proof. Note that the statement is equivalent to $\forall \epsilon \in]0,\infty[: \exists \delta \in]0,\infty[: f(B_A(a,\delta)) \subseteq B_B(b,\epsilon)$. Since the open balls form topological bases for the topologies of A and B, we know that for any open neighbourhood V of b in B there exists an $\epsilon \in]0,\infty[$ such that $b \in B_B(b,\epsilon) \subseteq V$ (recall Lemma (2.1.7)), and similarly for all open neighbourhoods U of a in A there exists a $\delta \in]0,\infty[$ such that $a \in B_A(a,\delta) \subseteq U$. □

Lemma 2.5.6

Let A, B $\boxed{\mathrm{T}}$, C $\boxed{d(\cdot,\cdot)}$, $a \in A$, $b \in B$, $c \in C$, and $f : A \times B \to C$.

Suppose

$$\lim_{(x,y)\to(a,b)} f(x,y) = c$$

and that for all $x \in A$ there exists a $g(x) \in C$ such that

$$\lim_{y\to b} f(x,y) = g(x),$$

this gives us a function $g : A \to C$.

Then for this function g we have

$$\lim_{x\to a} g(x) = c.$$

Proof. We follow [Dui2003]. Let $\epsilon \in]0,\infty[$ be given, then because $\lim_{(x,y)\to(a,b)} f(x,y) = c$ (by Lemma (2.5.5)) there exist open neighbourhoods U and V of a and b in A and B respectively such that for all $(x,y) \in U \times V$ we have $d_C(f(x,y),c) < \epsilon/2$. Let $x \in U$ and let $\delta \in]0,\infty[$ be arbitrary. Then because $\lim_{y\to b} f(x,y) = g(x)$

there exists an open neighbourhood V_δ of b in B such that for all $y \in V_\delta$, $d_C(f(x,y), g(x)) < \delta$. As for all $y \in V \cap V_\delta$, $(x,y) \in U \times V$, we find

$$d_C(g(x), c) \leq d_C(g(x), f(x,y)) + d_C(f(x,y), c) < \delta + \epsilon/2.$$

Since this is true for all $\delta \in]0,\infty[$ we find that necessarily

$$d_C(g(x), c) \leq 0 + \epsilon/2.$$

Therefore, for all $x \in U$ we have

$$d_C(g(x), c) \leq \epsilon/2 < \epsilon.$$

Hence (Lemma (2.5.5)), $\lim_{x \to a} g(x) = c$, as desired. $\square$

Corollary 2.5.7: Exchanging of limits

Let A, B, C, $a \in A$, $b \in B$, $c \in C$, and $f : A \times B \to C$.

If $\lim_{(x,y)\to(a,b)} f(x,y)$ exists and there exists an open neighbourhood $U \times V$ of (a,b) in $A \times B$ such that for all $(x,y) \in U \times V$ the limits

$$\lim_{x' \to a} f(x', y), \qquad \lim_{y' \to b} f(x, y')$$

exist, then

$$\lim_{x \to a}\left(\lim_{y \to b} f(x,y)\right) = \lim_{y \to b}\left(\lim_{x \to a} f(x,y)\right) = \lim_{(x,y)\to(a,b)} f(x,y).$$

Proof. Apply Lemma (2.5.6) to x and y separately for the metric spaces U and V and use that the limits exist in these metric spaces if and only if they exist in A and B, because $U \times V$ is an open neighbourhood of (a,b) in $A \times B$. $\square$

Definition 2.5.8

Let A be a (pseudo)metric space.

Then for any nonempty subset $B \subseteq A$ we define for all $a \in A$ the *distance from a to B* as

$$d(a, B) := \inf\{d(a, a_1) \in \mathbb{R} | a_1 \in B\} \geq 0.$$

Lemma 2.5.9

Let A be a (pseudo)metric space and $B \subseteq A$ a nonempty subset.

Then both the metric $d : A \times A \to \mathbb{R}$ and the distance function $d(\cdot, B) : A \to \mathbb{R} : a \mapsto d(a, B)$ are continuous.

Furthermore, $d(\cdot, B)^{-1}(\{0\}) = \overline{B}$.

Proof. First the metric. Let $(a_1, a_2) \in A \times A$, and $\epsilon \in]0,\infty[$ be given. Choose $U := B_A(a_1, \epsilon/2) \times B_A(a_2, \epsilon/2) \subseteq A \times A$ which is an open neighbourhood of (a_1, a_2) in $A \times A$. Let $(a_3, a_4) \in U$, then $d(a_3, a_4) \leq d(a_3, a_1) + d(a_1, a_4) \leq d(a_3, a_1) + d(a_1, a_2) + d(a_2, a_4) < \epsilon/2 + d(a_1, a_2) + \epsilon/2$, so $d(a_3, a_4) - d(a_1, a_2) < \epsilon$. On the other hand $d(a_1, a_2) \leq d(a_1, a_3) + d(a_3, a_2) \leq d(a_1, a_3) + d(a_3, a_4) + d(a_4, a_2) < \epsilon/2 + d(a_3, a_4) + \epsilon/2$, so $d(a_1, a_2) - d(a_3, a_4) < \epsilon$. So for all $(a_3, a_4) \in U$ we have $|d(a_3, a_4) - d(a_1, a_2)| < \epsilon$, hence $d(U) \subseteq]d(a_1, a_2) - \epsilon, d(a_1, a_2) + \epsilon[$. Since this is true for all $a_1, a_2 \in A$, $\epsilon \in]0,\infty[$, we have d .

Fix $B \subseteq A$ nonempty, then for any $a_1, a_2 \in A$ and $b \in B$ we have $d(a_1, B) \leq d(a_1, b) \leq d(a_1, a_2) + d(a_2, b)$, so $d(a_1, B) - d(a_1, a_2) \leq d(a_2, B)$ and hence

$d(a_1, B) - d(a_2, B) \leq d(a_1, a_2)$. Similarly $d(a_2, B) - d(a_1, B) \leq d(a_1, a_2)$, so $|d(a_1, B) - d(a_1, B)| \leq d(a_1, a_2)$. This gives continuity of $a \mapsto d(a, B)$.

Note that $\{0\} \subseteq \mathbb{R}$ is closed and as $d(\cdot, B)$ C, $d(\cdot, B)^{-1}(\{0\}) \subseteq A$ is closed as well by Lemma (2.1.14). For all $b \in B$ we have $0 \leq d(b, B) \leq d(b, b) = 0$, so $B \subseteq d(\cdot, B)^{-1}(\{0\})$. Therefore $\overline{B} \subseteq d(\cdot, B)^{-1}(\{0\})$. Let $a \in d(\cdot, B)^{-1}(\{0\})$ and U any open neighbourhood of a in A. Then there exists a $\delta \in]0, \infty[$ such that $a \in B_A(a, \delta) \subseteq U$. Since $d(a, B) = \inf\{d(a, b) | b \in B\} = 0 < \delta$ there exists some $b_\delta \in B$ such that $d(a, b_\delta) < \delta$. But then $b_\delta \in B_A(a, \delta) \subseteq U$ and therefore $B \cap U \neq \emptyset$. Because this is true for all open neighbourhoods U of a in A, we see (Lemma (2.1.3)) that $a \in \overline{B}$. Therefore $d(\cdot, B)^{-1}(\{0\}) \subseteq \overline{B}$. Because of this $\overline{B} = d(\cdot, B)^{-1}(\{0\})$. □

Theorem 2.5.10

Let A T.

Then A d(·,·) $\Rightarrow A$ T6 1C.

Proof. Let $a \in A$ be fixed. Consider the countable collection

$$\mathcal{A} := \left\{ B_A\left(a, \frac{1}{k}\right) \mid k \in \mathbb{N}, k \geq 1 \right\}.$$

Each element of $\mathcal{A}$ is clearly an open neighbourhood of a in A and for any open neighbourhood U of a in A, we have by definition of the topology that there exists an $\epsilon \in]0, \infty[$ such that $a \in B_A(a, \epsilon) \subseteq U$. Therefore, for $k \geq \lceil \frac{1}{\epsilon} \rceil \in \mathbb{N}$ we see that $a \in B_A(a, \frac{1}{k}) \subseteq B_A(a, \epsilon) \subseteq U$. Because of this A 1C.

Let $a_1, a_2 \in A$ and suppose that $a_1 \neq a_2$. Then $d(a_1, a_2) > 0$, choose $\epsilon = d(a_1, a_2)/2 > 0$, then $B_A(a_1, \epsilon)$ and $B_A(a_2, \epsilon)$ are two disjoint open neighbourhoods of a_1 resp. a_2 in A. Therefore A T2.

Let $B, C \subseteq A$ be disjoint and closed. Choose $f : A \to \mathbb{R}$ defined by $f(a) = d(a, B)/(d(a, B) + d(a, C))$. From Lemma (2.5.9) we know that $d(\cdot, B)^{-1}(\{0\}) = \overline{B} = B$, because B is closed. Therefore, for all $a \in A$ we have that $d(a, B) = 0$ if and only if $a \in B$ (and similarly for C). Because of this $d(a, B) + d(a, C) \leq 0$ if and only if $d(a, B) = d(a, C) = 0$ if and only if $a \in B \cap C = \emptyset$, which is impossible. So $d(a, B) + d(a, C) > 0$ for all $a \in A$. Therefore (together with continuity of d from Lemma (2.5.9)), f C. Let $a \in A$, then $f(a) = 0$ if and only if $d(a, B) = 0$ if and only if $a \in B$, so $f^{-1}(\{0\}) = B$. Furthermore, $f(a) = 1$ if and only if $d(a, B) = d(a, B) + d(a, C)$ if and only if $d(a, C) = 0$ if and only if $a \in C$, so $f^{-1}(\{1\}) = C$. So $f : A \to \mathbb{R}$ C and $f^{-1}(\{0\}) = B$, $f^{-1}(\{1\}) = C$. Since such a function exists for all disjoint, closed $B, C \subseteq A$ we see that A is perfectly normal and (we already saw A T2) therefore A T6. □

Definition 2.5.11: Cauchy sequence

Let A d(·,·) and $x : \mathbb{N} \to A$ a sequence.

Then we call x a *Cauchy sequence in* A if for all $\epsilon \in]0, \infty[$ there exists a $k \in \mathbb{N}$ such that for all $l, m \geq k$ we have $d(x_l, x_m) < \epsilon$.

Lemma 2.5.12

Let A d(·,·) and $x : \mathbb{N} \to A$ a sequence.

If x is convergent in A, then x is a Cauchy sequence in A.

Proof. Suppose x is convergent in A, then there exists an $a \in A$ such that $\lim_{k\to\infty} x_k = a$. Let $\epsilon \in]0, \infty[$, then because $B_A(a, \epsilon/2)$ is an open neighbourhood of a in A we have that there exists a $k \in \mathbb{N}$ such that for all $l \geq k$

we have $d(a, x_l) < \epsilon/2$. But then for all $l, m \geq k$ we have that $d(x_l, x_m) \leq d(x_l, a) + d(a, x_m) < \epsilon/2 + \epsilon/2 = \epsilon$. Therefore x is Cauchy. $\square$

Example 2.5.13

The converse of Lemma (2.5.12) is not true: take any sequence in $\mathbb{Q}$ approximating $\sqrt{2}$ with fractions (i.e. a decimal expansion where x_k is the k-decimal approximation: 1, 1.4, 1.41, 1.414, ...), then this sequence is Cauchy, but it has no limit in $\mathbb{Q}$ (as $\sqrt{2} \notin \mathbb{Q}$).

Definition 2.5.14: Completeness

Let A d(·,·).

We call A *complete* if every Cauchy sequence in A is convergent in A.

Example 2.5.15

Completeness is not a topological property: consider $]-1,1[$ and $\mathbb{R}$ both equipped with the usual absolute value metric $|\cdot|$. Then $\mathbb{R}$ is complete by construction, while for $]-1,1[$ we have that the Cauchy sequence $x : \mathbb{N} \to]-1,1[: k \mapsto 1 - \frac{1}{k+2}$ has no limit in $]-1,1[$ (the sequence converges to $1 \notin]-1,1[$ when viewed as a sequence in $\mathbb{R}$), hence $]-1,1[$ is not complete. However, $]-1,1[$ is T-isomorphic to $\mathbb{R}$ via $]-1,1[\to \mathbb{R} : x \mapsto \frac{x}{\sqrt{1-x^2}}$ with inverse $\mathbb{R} \to]-1,1[: x \mapsto \frac{x}{\sqrt{1+x^2}}$.

Therefore completeness is not preserved through T-isomorphism.

Definition 2.5.16: Baire space

Let A T.

Then we call A *Baire* if for any countable collection $B_1, B_2, \ldots$ of subsets of A where for all $k \in \mathbb{N}$, $B_k \subseteq A$ is closed and $\mathrm{int}(B_k) = \emptyset$ we have

$$\mathrm{int}\Big(\bigcup_{k\in\mathbb{N}} B_k\Big) = \emptyset.$$

Theorem 2.5.17: Baire category theorem

Let A d(·,·).

If A is complete, then A is Baire.

Proof. We follow [Mun2000]. Let $B_1, B_2, \ldots$ be a given countable collection of closed subsets of A with empty interiors. Let $U \subseteq A$ be any nonempty open set. As A d(·,·), we have by Theorem (2.5.10) that A T3. Since $\mathrm{int}(B_1) = \emptyset$ and $U \neq \emptyset$, there exists an $a_1 \in U \setminus B_1$. Now A T3, $a_1 \in U$, $a_1 \notin B_1$, and $B_1 \subseteq A$ is closed. Hence there exists a $\delta_1 \in]0,1[$ such that $U_1 := B_A(a_1, \delta_1) \subseteq \overline{U_1} \subseteq U$, $\overline{U_1} \cap B_1 = \emptyset$.

Suppose that we have constructed a sequence $(a_1, \delta_1), \ldots, (a_k, \delta_k) \in A \times]0, \infty[$, such that for each $1 \leq l \leq k$ we have $\delta_l \leq \frac{1}{l}$, $a_l \in U_l = B_A(a_l, \delta_l)$, $U_l \subseteq U_{l-1}$ (where $U_0 := U$), $\overline{U_l} \cap B_l = \emptyset$. Then as $\mathrm{int}(B_{k+1}) = \emptyset$, there exists an $a_{k+1} \in U_k \setminus B_{k+1}$ and (A T3) therefore there exists a $\delta_{k+1} \in]0, \frac{1}{k+1}[$ with $U_{k+1} := B_A(a_{k+1}, \delta_{k+1}) \subseteq U_k$ and $\overline{U_{k+1}} \cap B_{k+1} = \emptyset$.

Using induction this permits us to construct a sequence $\mathbb{N} \to A \times]0,1[: k \mapsto (a_k, \delta_k)$ with (here $U_k := B_A(a_k, \delta_k)$)

$$U = U_0 \supseteq U_1 \supseteq U_2 \supseteq \ldots, \qquad \forall k \in \mathbb{N} : \delta_k \leq \frac{1}{k}, \overline{U_k} \cap B_k = \emptyset.$$

By construction, for any given $\epsilon \in]0,\infty[$, take $k = \lceil \frac{1}{\epsilon} \rceil \in \mathbb{N}$, then for all $l, m \in \mathbb{N}$ with $k \leq l \leq m$ we have $a_m \in U_m \subseteq U_l \subseteq U_k$, $a_l \in U_l \subseteq U_k$, so $d(a_l, a_m) < \delta_k \leq \frac{1}{k} \leq \epsilon$. But this makes the sequence $\mathbb{N} \to A : k \mapsto a_k$ Cauchy and since A is complete, there exists an $a \in A$ such that $\lim_{k\to\infty} a_k = a$. As for all $k \in \mathbb{N}$, $a_l \in U_l \subseteq U_k$ for all $l \geq k$, necessarily $a \in \overline{U_k}$ (Lemma (2.3.5)). Since this is true for all $k \in \mathbb{N}$, we have $a \in \bigcap_{k\in\mathbb{N}} \overline{U_k}$. But this means that for all $k \in \mathbb{N}$, $a \notin B_k$, so $a \notin \bigcup_{k\in\mathbb{N}} B_k$. On the other hand $a \in \overline{U_1} \subseteq U$, so $a \in U \setminus \bigcup_{k\in\mathbb{N}} B_k$. Therefore U cannot be a subset of $\bigcup_{k\in\mathbb{N}} B_k$.

As this is true for all nonempty open sets $U \subseteq A$, the interior of $\bigcup_{k\in\mathbb{N}} B_k$ must be empty and hence A is Baire. □

Lemma 2.5.18

Let A [d(·,·)] and complete, and $B \subseteq A$ any subset.

Then B is closed if and only if B is complete (considered as a metric space with the restriction of A's metric to $B \times B$).

Proof. Suppose $B \subseteq A$ is closed. Let $x : \mathbb{N} \to B$ be any Cauchy sequence in B. Then x is a Cauchy sequence in A (as B has the metric from A restricted to $B\times B$), so because A is complete, there exists an $a \in A$ such that $\lim_{k\to\infty} x_k = a$. Hence by Lemma (2.3.5) $a = \lim_{k\to\infty} x_k \in \overline{B} = B$, as B is closed. So x converges in B. Hence B is complete.

Suppose $B \subseteq A$ is complete. Let $b \in \overline{B}$, then by Lemma (2.3.5) (A, B [1C] by Theorem (2.5.10)) there is a sequence $x : \mathbb{N} \to B$ such that $\lim_{k\to\infty} x_k = b$ in A. By Lemma (2.5.12), x is a Cauchy sequence in A. Since $x_k \in B$ for all $k \in \mathbb{N}$ and the metric on B is the restriction of the metric on A, x is a Cauchy sequence in B. As B is complete, there exists a $b' \in B$ such that $\lim_{k\to\infty} x_k = b'$. By Theorem (2.5.10), A, B [T2] so (Lemma (2.2.6)) necessarily $b = b' \in B$. Therefore $\overline{B} \subseteq B$ and hence B is closed. □

Definition 2.5.19: Lipschitz continuity

Let A, B [d(·,·)] and $\delta \in]0,\infty[$.

Then we call a map $f : A \to B$ δ-*Lipschitz continuous* (denoted by f δ-(C)) if

$$\forall a_1, a_2 \in A : d_B(f(a_1), f(a_2)) \leq \delta\, d_A(a_1, a_2).$$

Lemma 2.5.20

Let A, B [d(·,·)] and $f : A \to B$.

If f δ-(C) then f (C).

Proof. Fix any $a \in A$ and use Lemma (2.5.5). Let $\epsilon \in]0,\infty[$ be given, choose $\gamma = \epsilon/\delta > 0$, then for any $a_1 \in A$ satisfying $d_A(a, a_1) < \gamma$ we have (δ-(C)) $d_B(f(a), f(a_1)) \leq \delta\, d_B(a, a_1) < \delta\, \gamma = \epsilon$. Therefore $\lim_{x\to a} f(x) = f(a)$ and f (C) a. Since this is true for all $a \in A$, f (C). □

Theorem 2.5.21: Fixed point theorem

Let A [d(·,·)] complete and $f : A \to A$.

If f δ-(C) for $\delta \in]0, 1[$, then there exists a unique $a \in A$ such that $f(a) = a$.

Proof. We follow [DK2004I].

Suppose the conditions of the theorem are satisfied.

Let $a_1 \in A$ be arbitrary and construct the sequence $x : \mathbb{N} \to A$ by

$$x_k := (\underbrace{f \circ f \circ \ldots \circ f}_{k \text{ times}})(a_1),$$

such that $x_{k+1} = f(x_k)$ for all $k \in \mathbb{N}$.

Then for $k \in \mathbb{N}$, $k \geq 2$ we have

$$d(x_k, x_{k+1}) = d(f(x_{k-1}), f(x_k)) \leq \delta\, d(x_{k-1}, x_k) \leq \ldots \leq \delta^{k-1}\, d(x_1, x_2).$$

So for $k, l \in \mathbb{N}$,

$$\begin{aligned} d(x_k, x_{k+l}) &\leq d(x_k, x_{k+1}) + d(x_{k+1}, x_{k+2}) + \ldots + d(x_{k+l-1}, x_{k+l}) \\ &\leq \delta^{k-1}\, d(x_1, x_2) + \delta^{k}\, d(x_1, x_2) + \ldots + \delta^{k+l-2}\, d(x_1, x_2) \\ &= \delta^{k-1}(1 + \delta + \ldots + \delta^{l-1})\, d(x_1, x_2) \\ &\leq \frac{\delta^{k-1}}{1-\delta}\, d(x_1, x_2). \end{aligned}$$

As $\delta \in]0,1[$, this term can be made arbitrarily small when we increase k. Therefore the sequence x is Cauchy and because A is complete, this means that there exists an $a \in A$ such that $\lim_{k\to\infty} x_k = a$. By the chain rule for limits we have (use Lemma (2.5.20) for continuity of f)

$$f(a) = f\Big(\lim_{k\to\infty} x_k\Big) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} x_{k+1} = a.$$

Therefore $f(a) = a$, so a is a fixed point of f.

Suppose $a_1 \in A$ satisfies $f(a_1) = a_1$, then $d(a, a_1) = d(f(a), f(a_1)) \leq \delta\, d(a, a_1)$, so $(1-\delta)\, d(a, a_1) \leq 0$ and since $1 - \delta > 0$ and $d(a_0, b_0) \geq 0$ this implies that $d(a, a_1) = 0$ and hence $a = a_1$. Therefore a is unique. □

Definition 2.5.22: Uniform continuity

Let A, B.

Then we call a map $f : A \to B$ *uniformly continuous* if

$$\forall \epsilon \in]0, \infty[: \exists \delta \in]0, \infty[: \forall a_1, a_2 \in A : (d_A(a_1, a_2) < \delta \to d_B(f(a_1), f(a_2)) < \epsilon).$$

Compare Definition (2.5.22) with a function $f : A \to B$ being :

$$\forall \epsilon \in]0, \infty[: \forall a_1 \in A : \exists \delta \in]0, \infty[: \forall a_2 \in A : (d_A(a_1, a_2) < \delta \to d_B(f(a_1), f(a_2)) < \epsilon);$$

with uniform continuity we can pick a *single* δ for all a_1, while for ordinary continuity, this δ may vary wildly as we vary a_1.

Theorem 2.5.23: Uniform continuity

Let A, B and $f : A \to B$.

If A, then f is uniformly continuous.

Proof. We follow [DK2004I].

Suppose f is not uniformly continuous. Then there exists an $\epsilon \in]0, \infty[$ such that for all $\delta \in]0, \infty[$ there exist $a_1, a_2 \in A$ such that $d_A(a_1, a_2) < \delta$, while $d_B(f(a_1), f(a_2)) \geq \epsilon$. In particular for each $\delta = 1/k$, $k \in \mathbb{N}$ there exist $a_k, a_k' \in A$ such that $d_A(a_k, a_k') < 1/k$ and $d_B(f(a_k), f(a_k')) \geq \epsilon$. Because A

1C (by Theorem (2.5.10)) and A Cpt, by Lemma (2.4.6) the sequences $k \mapsto a_k$ and $k \mapsto a'_k$ have convergent subsequences x and x' with limits $a, a' \in A$ respectively. Because $d_A(a_k, a'_k) < 1/k$ for all $k \in \mathbb{N}$ we see that necessarily $a = a'$ and hence, because f C,

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f(x'_k) = f(a).$$

Hence (Lemma (2.5.9)) $\lim_{k\to\infty} d_B(f(x_k), f(x'_k)) = d_B(f(a), f(a)) = 0$, contradicting the fact that $d_B(f(a_k), f(a'_k)) \geq \epsilon$ for all $k \in \mathbb{N}$. So we reach a contradiction: f must be uniformly continuous. $\square$

Chapter 3

Algebra

Algebra deals with the quantitative study of counting and symmetries.

3.1 Groups

Definition 3.1.1: Group (G)
Let A be a set.

A *group structure on* A is an element $e \in A$, called the *identity*, together with two maps

$$A \times A \to A : (a_1, a_2) \mapsto a_1\, a_2, \qquad A \to A : a_1 \mapsto a_1^{-1},$$

called *multiplication* and *inversion* respectively, that satisfy for all $a_1, a_2, a_3 \in A$,

- $e\, a_1 = a_1\, e = a_1$,
- $(a_1\, a_2)\, a_3 = a_1\, (a_2\, a_3)$,
- $a_1\, a_1^{-1} = a_1^{-1}\, a_1 = e$.

We call a group structure on A *Abelian* or *commutative* if for all $a_1, a_2 \in A$ we have $a_1\, a_2 = a_2\, a_1$. In this case we usually denote e by 0 (called *zero*), $a_1\, a_2$ by $a_1 + a_2$ (called *addition*), a_1^{-1} by $-a_1$ (called *negation*), and $a_1 + (-a_2)$ by $a_1 - a_2$.

A *group* A (denoted by A G) is a set A together with a group structure.

With all objects that have algebraic properties we will use the notation $B \leq A$ to indicate that $B \subseteq A$ is a subset and that the restriction of all algebraic maps (addition, multiplication, ...) to B, makes B an algebraic object of the same type as A.

Example 3.1.2
Let $A = \mathbb{R} \setminus \{0\}$ considered as an Abelian group with multiplication and division on $\mathbb{R}$. Then $B := \{-1, +1\} \leq A$ since B G with respect to (the restrictions of) multiplication and division on $\mathbb{R}$ to B.

Lemma 3.1.3
Let A $\boxed{\text{G}}$.

Then the identity and inverses of elements of A are unique, and for all $a_1, a_2 \in A$ we have $(a_1\, a_2)^{-1} = a_2^{-1}\, a_1^{-1}$.

Proof. Suppose there exists an $a \in A$ such that for all $a_1 \in A$ we have $a\, a_1 = a_1\, a = a_1$. Then in particular $e = a\, e = a$, so $a = e$. Therefore the identity of A is unique.

Let $a \in A$ be arbitrary and suppose $a_1, a_2 \in A$ satisfy $a\, a_1 = a_1\, a = e$ and $a\, a_2 = a_2\, a = e$. Then $a_1 = e\, a_1 = (a_2\, a)\, a_1 = a_2\, (a\, a_1) = a_2\, e = a_2$, so $a_1 = a_2$ and the inverse of a is unique.

We have that $(a_1\, a_2)(a_2^{-1}\, a_1^{-1}) = a_1\, (a_2\, a_2^{-1})\, a_1^{-1} = a_1\, e\, a_1^{-1} = a_1\, a_1^{-1} = e$ and similarly $(a_2^{-1}\, a_1^{-1})(a_1\, a_2) = e$, so by uniqueness of inverses $(a_1\, a_2)^{-1} = a_2^{-1}\, a_1^{-1}$. □

Example 3.1.4: Group of permutations
Let $k \in \mathbb{N}$, then the *group of permutations of* $\{1, \ldots, k\}$ is defined to be the set

$$S^k := \{\pi : \{1, \ldots, k\} \to \{1, \ldots, k\} \mid \pi \text{ bijective } \}.$$

together with identity $\mathrm{id}_{\{1,\ldots k\}}$, multiplication $(\pi_1, \pi_2) \mapsto \pi_1\, \pi_2 := \pi_1 \circ \pi_2$, and inversion $\pi \mapsto \pi^{-1}$ (the inverse function associated with the bijective function π).

Definition 3.1.5: Morphisms of groups
Let A, B $\boxed{\text{G}}$.

Then all maps $f : A \to B$ that satisfy for all $a_1, a_2 \in A$,

$$f(a_1\, a_2) = f(a_1)\, f(a_2)$$

are *group morphisms between* A *and* B (denoted by f $\boxed{\text{G}}$-morphism). For all group morphisms f, the *kernel of* f is defined as

$$\ker f := \{a \in A \mid f(a) = e_B\}.$$

The *identity morphism of* A is the map

$$\mathrm{id}_A : A \to A : a \mapsto a.$$

Lemma 3.1.6
Let A, B $\boxed{\text{G}}$ and $f : A \to B$ a $\boxed{\text{G}}$-morphism.

Then

- $f(e_A) = e_B$,
- for all $a \in A$ we have $f(a^{-1}) = f(a)^{-1}$,
- $\{e_A\} \leq \ker f \leq A$ and $\ker f = \{e_A\}$ if and only if f is injective,
- f is a $\boxed{\text{G}}$-isomorphism if and only if f is bijective.

Proof.
- Let $b \in f(A) \subseteq B$ ($f(A) \supseteq f(\{e_A\}) \neq \emptyset$) then there exists an $a \in A$ such that $f(a) = b$. Now $f(e_A)\, b = f(e_A)\, f(a) = f(e_A\, a) = f(a) = b$, so $f(e_A) = f(e_A)\, e_B = f(e_A)\, (b\, b^{-1}) = (f(e_A)\, b)\, b^{-1} = b\, b^{-1} = e_B$, which shows the first statement.

- Note that $f(a^{-1})\,f(a) = f(a^{-1}\,a) = f(e_A) = e_B$ and similarly $f(a)\,f(a^{-1}) = e_B$, so by uniqueness $f(a^{-1}) = f(a)^{-1}$.

- Since $f(e_A) = e_B$, $e_A \in \ker f$ and we see immediately that $\{e_A\} \leq A$. Suppose $a_1, a_2 \in \ker f$, then $f(a_1\,a_2) = f(a_1)\,f(a_2) = e_B\,e_B = e_B$, so $a_1\,a_2 \in \ker f$, hence multiplication restricts to $\ker f \times \ker f \to \ker f$. Suppose $a_1 \in \ker f$, then $f(a_1^{-1}) = f(a_1)^{-1} = e_B^{-1} = e_B$, so inversion restricts to $\ker f \to \ker f$. Therefore $\ker f \leq A$ and we obtain $\{e_A\} \leq \ker f \leq A$.

  Suppose f is injective. Let $a_1 \in \ker f$, then $f(a_1) = e_B = f(e_A)$, so $a_1 = e_A$. Therefore $\ker f \subseteq \{e_A\}$ and hence $\ker f = \{e_A\}$.

  Suppose $\ker f = \{e_A\}$. Let $a_1, a_2 \in A$ with $f(a_1) = f(a_2)$, then $e_B = f(a_1)^{-1}\,f(a_1) = f(a_2)^{-1}\,f(a_1) = f(a_2^{-1}\,a_1)$, so $a_2^{-1}\,a_1 \in \ker f = \{e_A\}$, hence $a_2^{-1}\,a_1 = e_A$, so $a_1 = a_2\,a_2^{-1}\,a_1 = a_2 e_A = a_2$. This makes f injective.

- Suppose f is an isomorphism, then there exists a morphism of groups $g : B \to A$ such that $f \circ g = \mathrm{id}_B$, $g \circ f = \mathrm{id}_A$, so f is a bijection with inverse g.

  Suppose conversely that f is a bijection, let $g : B \to A$ denote the inverse of f, then $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. We need to check that g is a $\boxed{\mathbf{G}}$-morphism. Let $b_1, b_2 \in B$, then $g(b_1\,b_2) = g(f(g(b_1))\,f(g(b_2))) = g(f(g(b_1)\,g(b_2))) = g(b_1)\,g(b_2)$ because f is a $\boxed{\mathbf{G}}$-morphism, so g is a $\boxed{\mathbf{G}}$-morphism and therefore f is a $\boxed{\mathbf{G}}$-isomorphism.

$\square$

Definition 3.1.7: Normal subgroup
Let A $\boxed{\mathbf{G}}$.

Then any subset $B \subseteq A$ is called a *normal subgroup of* A if $B \leq A$ and for all $b \in B$ and $a \in A$ we have that $aba^{-1} \in B$.

Definition 3.1.8: Quotient group
Let A $\boxed{\mathbf{G}}$, $B \leq A$ a normal subgroup.

The *quotient group* A/B is defined as the set of equivalence classes

$$A/B := \{[a] \mid a \in A\},$$

with $[a] := a\,B := \{a\,b \in A | b \in B\}$, together with identity $[e] = B$, multiplication $[a_1][a_2] := [a_1\,a_2]$, and inversion $[a_1]^{-1} := [a_1^{-1}]$.

That this is a group can be verified using Definition (3.1.7). First note that for any $a \in A$, $b \in B$ we have $[a] = [a\,b] = [b\,a]$, since for any $b_1 \in B$, $b\,a\,b_1 = a\,((a^{-1}\,b\,a)\,b_1) \in [a]$. Suppose $[a_1] = [a_3]$ and $[a_2] = [a_4]$, then $[a_1\,a_2] = [a_3\,a_4]$: as $a_3 \in [a_1]$ there is a $b_3 \in B$ such that $a_3 = a_1\,b_3$ and similarly $a_4 = a_2\,b_4$ for some $b_4 \in B$, hence $[a_3\,a_4] = [(a_1\,b_3\,a_2)\,b_4] = [a_1\,b_3\,a_2] = [a_1\,a_2\,(a_2^{-1}\,b_3\,a_2)] = [a_1\,a_2]$. This makes multiplication well-defined. For inversion note that $[a_1] = [a_3]$ iff for some $b_3 \in B$, $a_3 = a_1\,b_3$ iff $a_1^{-1}\,a_3 \in B$ iff $a_3^{-1}\,(a_1^{-1})^{-1} \in B$ iff $[a_1^{-1}] = [a_3^{-1}]$, so inversion is well-defined. Now all the desired group properties follow from pulling them inside the brackets $[\ldots]$.

Lemma 3.1.9: Factorisation
Let A, B $\boxed{\mathbf{G}}$.

Then for any $f : A \to B$ $\boxed{\mathbf{G}}$-morphism,

- $\ker f \leq A$ is a normal subgroup,
- $\{e_B\} \leq f(A) \leq B$ and $f(A) = B$ if and only if f is surjective,
- there exists a unique, injective **G**-morphism $g : A/\ker f \to B$ such that $f(a) = g([a])$ for all $a \in A$,
- $f(A) \simeq A/\ker f$ are **G**-isomorphic.

Conversely, for any normal subgroup $C \leq A$ there exist a group D and **G**-morphism $f : A \to D$ such that $\ker f = C$.

Proof. Let $f : A \to B$ be a **G**-morphism.

- By Lemma (3.1.6) we know that $\ker f \leq A$. Let $a \in A$, $b \in \ker f$, then $f(a\,b\,a^{-1}) = f(a)\,f(b)\,f(a)^{-1} = f(a)\,e_A\,f(a)^{-1} = f(a)\,f(a)^{-1} = e_B$, so $a\,b\,a^{-1} \in \ker f$. Hence $\ker f$ is a normal subgroup.
- Let $b_1, b_2 \in f(A)$, then $b_1 = f(a_1)$, $b_2 = f(a_2)$ for $a_1, a_2 \in A$, hence $b_1\,b_2 = f(a_1)\,f(a_2) = f(a_1\,a_2) \in f(A)$. Furthermore by Lemma (3.1.6) $b_1^{-1} = f(a_1)^{-1} = f(a_1^{-1}) \in f(A)$ and $e_B = f(e_A) \in f(A)$. Hence $\{e_B\} \leq f(A) \leq B$. By definition $f(A) = B$ if and only if f is surjective.
- Choose $g : A/\ker f \to B$, $g([a]) := f(a)$. Then g is well-defined: if $[a_1] = [a_2]$, then $a_1\,a_2^{-1} \in \ker f$, so by Lemma (3.1.6) $e_B = f(a_1\,a_2^{-1}) = f(a_1)\,f(a_2)^{-1} = g([a_1])\,g([a_2])^{-1}$, so $g([a_1]) = g([a_2])$. Also $g([a_1]\,[a_2]) = g([a_1\,a_2]) = f(a_1\,a_2) = f(a_1)\,f(a_2) = g([a_1])\,g([a_2])$, so g is a **G**-morphism. It is clear that g is uniquely determined by definition. Suppose that $g([a_1]) = e_B$, then $f(a_1) = e_B$, so $a_1 \in \ker f$ and hence $[a_1] = [e_A]$, so $\ker g = \{[e_A]\}$ and by Lemma (3.1.6) g is injective.
- As g is injective and $g([a]) = f(a)$, $g(A/\ker g) = f(A)$, so $A/\ker f \to f(A) : [a] \mapsto g([a])$ is bijective. Hence by Lemma (3.1.6), it is a **G**-isomorphism.

Let $C \leq A$ be a normal subgroup. Choose $D := A/C$ and $f : A \to D$, $f(a) := [a]$. Then $f(a) = [e_A]$ iff $a \in C$, so $\ker f = C$. $\square$

Definition 3.1.10: Group action
Let A **G**, B a set.

Then an *action* f *of* A *on* B is a map $f : A \times B \to B : (a, b) \mapsto a \cdot b$ such that for all $a_1, a_2 \in A$, $b \in B$ we have

$$a_1 \cdot (a_2 \cdot b) = (a_1\,a_2) \cdot b, \qquad e_A \cdot b = b.$$

For any $b \in B$ we define *the orbit of* b to be

$$A \cdot b := \{a \cdot b \in B \mid a \in A\} \subseteq B$$

and the *stabilisers of* b to be

$$A_b := \{a \in A \mid a \cdot b = b\} \subseteq A.$$

We can check directly that defining $b_1 \sim b_2$ if and only if $A \cdot b_1 = A \cdot b_2$ is an equivalence relation, which shows that the orbits of a group action partition the set on which the group acts.

Theorem 3.1.11: Orbit-stabiliser theorem
Let A **G**, B a set.

Let f be an action of A on B, then for all $b \in B$ there is a bijection

$$A/A_b \simeq A \cdot b.$$

Proof. Let $b \in B$. Consider the morphism $g : A \to A \cdot b$ defined by $g(a) := a \cdot b$. The set $A \cdot b$ has a natural group structure $(a_1 \cdot b)(a_2 \cdot b) := (a_1\, a_2) \cdot b$, $(a_1 \cdot b)^{-1} := (a_1^{-1}) \cdot b$ with respect to which g is a group morphism. Now $a \cdot b = g(a) = e_A \cdot b = b$ if and only if $a \in A_b$, so $\ker g = A_b$. Furthermore, for any $(a_1 \cdot b) \in A \cdot b$, $g(a_1) = a_1 \cdot b$, so g is surjective. Therefore Lemma (3.1.9) gives us that the induced map $A/\ker g = A/A_b \to g(A) = A \cdot b$ is a **G**-isomorphism and hence a bijection. □

3.2 Rings

Definition 3.2.1: Ring
Let A be a set.

A *ring structure on* A consists of an Abelian group structure on A (so $0 \in A$, addition $(a_1, a_2) \mapsto a_1 + a_2$, negation $a_1 \mapsto -a_1$), together with an element $1 \in A$ (called *one*) and a map

$$A \times A \to A : (a_1, a_2) \mapsto a_1\, a_2$$

(called *multiplication*), that satisfy for all $a_1, a_2, a_3 \in A$ that

- $a_1\, 1 = 1\, a_1 = a_1$,
- $(a_1\, a_2)\, a_3 = a_1\, (a_2\, a_3)$,
- $a_1\, (a_2 + a_3) = (a_1\, a_2) + (a_1\, a_3)$,
- $(a_1 + a_2)\, a_3 = (a_1\, a_3) + (a_2\, a_3)$.

We call any $a \in A$ *invertible* with respect to this ring structure if there exists an $a^{-1} \in A$ such that $a\, a^{-1} = a^{-1}\, a = 1$. The collection of all invertible elements of A is denoted by

$$A^* := \{a \in A \mid a \text{ invertible}\}.$$

We say that this ring structure is *Abelian* or *commutative* if for all $a_1, a_2 \in A$ we have $a_1\, a_2 = a_2\, a_1$.

A *ring* A (denoted by A **R**) is a set A together with a ring structure.

Definition 3.2.2: Morphisms of rings
Let A, B **R**.

Then all maps $f : A \to B$ that are group morphisms of the Abelian group structures on A and B and satisfy for all $a_1, a_2 \in A$ that

$$f(a_1\, a_2) = f(a_1)\, f(a_2), \qquad f(1_A) = 1_B,$$

are *ring morphisms between A and B* (denoted by f **R**-morphism).

The *identity morphism of* A is the map

$$\mathrm{id}_A : A \to A : a \mapsto a.$$

So a map $f : A \to B$ between two rings is a ring morphism if and only if for all $a_1, a_2 \in A$ we have $f(a_1 + a_2) = f(a_1) + f(a_2)$, $f(a_1\, a_2) = f(a_1)\, f(a_2)$, and $f(1_A) = 1_B$.

Lemma 3.2.3: Solving equations in a ring
Let A **R**.

- For all $a \in A$, $0\, a = a\, 0 = 0$.
- For all $a \in A$ we have $-a = (-1)\, a$.
- $1 \in A^*$.
- $0 \in A^*$ if and only if $0 = 1$ if and only if $A = A^* = \{0\}$.
- 0 is uniquely determined and for all $a_1, a_2 \in A$, $-(a_1 + a_2) = -a_1 - a_2$ and $-a_1$ is unique.
- 1 is uniquely determined, inverses of invertible elements are unique, $(a_1\, a_2)^{-1} = a_2^{-1}\, a_1^{-1}$.
- If $a_1 \in A^*$, $a_1\, a_2 = a_1\, a_3$ if and only if $a_2 = a_3$ if and only if $a_2\, a_1 = a_3\, a_1$.
- $a_1 + a_2 = a_1 + a_3$ if and only if $a_2 = a_3$ if and only if $a_2 + a_1 = a_3 + a_1$.
- Suppose $A^* = A \setminus \{0\}$. If $a_1, a_2 \in A$ with $a_1\, a_2 = 0$, then $a_1 = 0$ or $a_2 = 0$.

Proof.
- Consider $0\, a = (0 + 0)\, a = (0\, a) + (0\, a)$, so $0 = (0\, a) - (0\, a) = ((0\, a) + (0\, a)) - (0\, a) = (0\, a) + ((0\, a) - (0\, a)) = (0\, a) + 0 = (0\, a)$. Similarly $a\, 0 = 0$.
- By using $0\, a = 0$ we find $a + ((-1)\, a) = (1\, a) + ((-1)\, a) = (1 + (-1))\, a = 0\, a = 0$ and similarly $((-1)\, a) + a = 0$, so by Lemma (3.1.3), $(-1)\, a = -a$.
- As $1\, 1 = 1\, 1 = 1$, $A \ni 1 = 1^{-1}$ exists and therefore $1 \in A^*$ is invertible.
- Suppose $0 \in A^*$, then $1 = 0\, 0^{-1} = 0$.

  Suppose $0 = 1$, then for any $a \in A$, $a = 1\, a = 0\, a = 0$ and $0 = 1 \in A^*$, so $A = A^* = \{0\}$.

  Suppose $A = A^* = \{0\}$, then $0 \in A^*$ directly.
- The fact that 0 and negations are uniquely determined follows directly from Lemma (3.1.3).
- The set A^* together with $1 \in A^*$, $(a_1, a_2) \mapsto a_1\, a_2$, and $a_1 \mapsto a_1^{-1}$ is a group. In particular, 1 is unique, all inverses of invertible elements are unique and $(a_1\, a_2)^{-1} = a_2^{-1}\, a_1^{-1}$ as follows from Lemma (3.1.3).
- Let $a_1 \in A^*$ and $a_2, a_3 \in A$. Suppose $a_1\, a_2 = a_1\, a_3$, then $a_2 = 1\, a_2 = (a_1^{-1}\, a_1)\, a_2 = a_1^{-1}\, (a_1\, a_2) = a_1^{-1}\, (a_1\, a_3) = (a_1^{-1}\, a_1)\, a_3 = 1\, a_3 = a_3$. Similarly if $a_2\, a_1 = a_3\, a_1$ then $a_2 = a_3$. The converse is immediate from multiplying by a_1 on the left or the right of $a_2 = a_3$.
- Results for addition follow in exactly the same way as for multiplication.

- Suppose $A^* = A \setminus \{0\}$ and $a_1, a_2 \in A$, $a_1 a_2 = 0$. Suppose $a_1 \neq 0$, then $a_1 \in A^*$, so $a_2 = a_1^{-1} a_1 a_2 = a_1^{-1} 0 = 0$. Similarly if $a_2 \neq 0$, then $a_1 = 0$. Hence either $a_1 = 0$ or $a_2 = 0$.

□

Definition 3.2.4: Ideal
Let A $\boxed{R}$.

Then we call a set B an *ideal of* A if $B \subseteq A$, $0 \in B$, and for all $a_1 \in A$, $b_1, b_2 \in B$ we have $a_1 b_1 - b_2 \in B$ and $b_1 a_1 - b_2 \in B$.

For any ideal $B \subseteq A$ we have that $B \neq A$ if and only if $B \cap A^* = \emptyset$ (if there exists a $b \in B \cap A^*$, then $1 = b^{-1} b - 0 \in B$, so $a = a1 - 0 \in B$ for all $a \in A$, so $A = B$).

Definition 3.2.5: Quotient ring
Let A $\boxed{R}$, and B an ideal of A.

Then the *quotient ring* A/B is defined as the set of equivalence classes

$$A/B := \{[a] \subseteq A | a \in A\},$$

with $[a] := \{a_1 \in A | a_1 - a \in B\}$, together with $0_{A/B} := [0_A]$, $1_{A/B} := [1_A]$, $[a_1] + [a_2] := [a_1 + a_2]$, $[a_1][a_2] := [a_1\, a_2]$.

That this indeed is a ring is easily verified from the fact that B is an ideal. Suppose $[a_1] = [a_3]$ and $[a_2] = [a_4]$, then $[a_1 + a_2] = [a_3 + a_4]$ ($a_3 \in [a_3] = [a_1]$, so $a_3 = a_1 + b_3$ for some $b_3 \in B$, similarly $a_4 = a_2 + b_4$, so $a_5 \in [a_1 + a_2]$ iff $B \ni a_5 - (a_1 + a_2) = a_5 - ((a_3 - b_3) + (a_4 - b_4)) = (a_5 - (a_3 + a_4)) + b_3 + b_4$ iff $a_5 - (a_3 + a_4) \in B$ iff $a_5 \in [a_3 + a_4]$), so addition is well defined. Similarly $a_5 \in [a_1 a_2]$ iff $B \ni a_5 - a_1\, a_2 = a_5 - (a_3 + b_3)(a_4 + b_4) = (a_5 - a_3\, a_4) - a_3\, b_4 - b_3\, a_4 - b_3\, b_4$ iff $a_5 - a_3\, a_4 \in B$ (as B is an ideal: $a_3\, b_4, b_3\, a_4 \in B$ since $b_3, b_4 \in B$) iff $a_5 \in [a_3\, a_4]$, so $[a_1\, a_2] = [a_3\, a_4]$ if $[a_1] = [a_3]$, $[a_2] = [a_4]$ which makes multiplication well-defined. All the requirements from Definition (3.2.1) are now satisfied by pulling every expression inside $[\ldots]$ and using the fact that they are satisfied by A.

Definition 3.2.6: Field
Let A $\boxed{R}$.

Then we call A a *field* (denoted by A $\boxed{F}$) if the ring structure on A is commutative and has $A^* = A \setminus \{0\}$.

By Lemma (3.2.3) we see that $A^* = A \setminus \{0\}$ implies that $0 \neq 1$.

3.3 Modules

Definition 3.3.1: Module
Let A $\boxed{R}$ and B a set.

Then an A-*module structure on* B consists of an Abelian group structure on B (so $0_B \in B$, $(b_1, b_2) \mapsto b_1 +_B b_2$, $b_1 \mapsto -b_1$), together with a map

$$A \times B \to B : (a, b) \mapsto a\, b.$$

(called *scalar multiplication*), satisfying for all $a_1, a_2 \in A$ and $b_1, b_2 \in B$ that

- $(a_1 +_A a_2)\, b_1 = (a_1\, b_1) +_B (a_2\, b_1)$,
- $a_1\, (b_1 +_B b_2) = (a_1\, b_1) +_B (a_1\, b_2)$,
- $(a_1\, a_2)\, b_1 = a_1\, (a_2\, b_1)$,
- $1_A\, b_1 = b_1$.

An A-*module* B (denoted by B $\boxed{\mathbf{M}}$ $/A$) is a set B together with an A-module structure.

In exactly the same way as for rings (Lemma (3.2.3)) we can show that for all $b \in B$ we have $-b = (-1_A)\, b$ and $0_A\, b = 0_B$.

Example 3.3.2: $\boxed{\mathbf{R}} \Rightarrow \boxed{\mathbf{M}}$
Let A $\boxed{\mathbf{R}}$, then A $\boxed{\mathbf{M}}$ $/A$ with respect to its own addition and multiplication as module operations.

Definition 3.3.3: A-linearity ($\textcircled{l}$ $/A$)
Let A $\boxed{\mathbf{R}}$, B, C $\boxed{\mathbf{M}}$ $/A$.
Then a map $f : B \to C$ is called A-*linear* (denoted by f $\textcircled{l}$ $/A$) if for all $a_1 \in A$ and $b_1, b_2 \in B$ we have

$$f(a_1\, b_1 +_B b_2) = a_1\, f(b_1) +_C f(b_2).$$

For an A-linear map $f : B \to C$ we define the *kernel of* f to be

$$\ker f := \{b \in B \mid f(b) = 0_C\}.$$

As with Lemma (3.1.6) we have (by considering f as a $\boxed{\mathbf{G}}$-morphism between the Abelian group structures of B and C) that f is injective if and only if $\ker f = \{0_B\}$.
From this lemma we also see that for any $f : B \to C$ $\textcircled{l}$ $/A$, $\{0_B\} \leq \ker f \leq B$ and $\{0_C\} \leq f(B) \leq C$.

Definition 3.3.4: Morphisms of A-modules
Let A $\boxed{\mathbf{R}}$, B, C $\boxed{\mathbf{M}}$ $/A$.
Then all maps $f : B \to C$ $\textcircled{l}$ $/A$ are *A-module morphisms between B and C*.
The *identity morphism of B* is the map

$$\mathrm{id}_B : B \to B : b \mapsto b.$$

Looking at Lemma (3.1.6) we also see that any $f : B \to C$ $\textcircled{l}$ $/A$ that is bijective, is in fact a $\boxed{\mathbf{M}}$-isomorphism.

Definition 3.3.5: Multilinearity (k-$\textcircled{l}$)
Let $k \in \mathbb{N}$, A $\boxed{\mathbf{R}}$, $B_1, \ldots, B_k$, C $\boxed{\mathbf{M}}$ $/A$.
Then we say that a map $f : B_1 \times \ldots \times B_k \to C$ is k-*multilinear over* A (denoted by f k-$\textcircled{l}$ $/A$) if for all $b_1 \in B_1, \ldots, b_k \in B_k$ and $1 \leq l \leq k$ the map

$$B_l \to C : b \mapsto f(b_1, \ldots, b_{l-1}, b, b_{l+1}, \ldots, b_k)$$

is $\textcircled{l}$ $/A$.

Definition 3.3.6: Quotient module

Let A $\boxed{\mathbf{R}}$, B $\boxed{\mathbf{M}}/A$, $C \leq B$.

Then the *quotient module* B/C is defined as the set of equivalence classes

$$B/C := \{[b] \subseteq B \mid b \in B\},$$

where $[b] := \{b_1 \in B | b_1 - b \in C\}$, together with $0_{B/C} := [0_B]$, $[b_1] + [b_2] := [b_1 + b_2]$, $a[b_1] := [a\, b_1]$.

That B/C is indeed $\boxed{\mathbf{M}}/A$ is verified in the same way as for Definition (3.2.5) (gives well-definedness of addition), together with the fact that if $[b_1] = [b_2]$, then $b_2 = b_1 + c_2$ for some $c_2 \in C$, so $b_3 \in [a\, b_1]$ iff $C \ni b_3 - a\, b_1 = b_3 - a\,(b_2 - c_2) = (b_3 - a\, b_2) - a\, c_3$ iff $b_3 - a\, b_2 \in C$ (as $a\, c_3 \in C$ because $C \leq B$), so $[a\, b_1] = [a\, b_2]$ and scalar multiplication is well-defined. Again all requirements of Definition (3.3.1) follow from pulling them inside $[\ldots]$.

Definition 3.3.7: Direct product

Let A $\boxed{\mathbf{R}}$, $\{B_i | i \in I\}$ with B_i $\boxed{\mathbf{M}}/A$ for all $i \in I$.

Then the *direct product of* $\{B_i | i \in I\}$ is defined as the set

$$\prod_{i \in I} B_i = \Big\{ g : I \to \bigcup_{i \in I} B_i \mid \forall i \in I : g(i) \in B_i \Big\},$$

together with zero $i \mapsto 0_{B_i}$, addition $(g_1 + g_2)(i) := g_1(i) +_{B_i} g_2(i)$, and scalar multiplication $(a\, g)(i) := a \cdot_{B_i} g(i)$.

Note that $\prod_{i \in I} B_i$ $\boxed{\mathbf{M}}/A$.

Definition 3.3.8: Direct sum

Let A $\boxed{\mathbf{R}}$, $\{B_i | i \in I\}$ with B_i $\boxed{\mathbf{M}}/A$ for all $i \in I$.

Then the *direct sum of* $\{B_i | i \in I\}$ is defined as the set

$$\bigoplus_{i \in I} B_i := \Big\{ g \in \prod_{i \in I} B_i \mid \{i \in I | g(i) \neq 0_{B_i}\} \text{ is finite } \Big\},$$

together with zero $i \mapsto 0_{B_i}$, addition $(g_1 + g_2)(i) := g_1(i) +_{B_i} g_2(i)$, and scalar multiplication $(a\, g)(i) := a \cdot_{B_i} g(i)$.

Note that $\bigoplus_{i \in I} B_i$ $\boxed{\mathbf{M}}/A$.

Lemma 3.3.9: Direct product and sum properties

Let A $\boxed{\mathbf{R}}$, $\{B_i | i \in I\}$ with B_i $\boxed{\mathbf{M}}/A$ for all $i \in I$.

- We always have that
$$\bigoplus_{i \in I} B_i \leq \prod_{i \in I} B_i.$$
- If $I = \{i_1, \ldots, i_k\}$ is finite, then as $\boxed{\mathbf{M}}/A$
$$\bigoplus_{i \in I} B_i \simeq \prod_{i \in I} B_i \simeq B_{i_1} \times \ldots \times B_{i_k},$$
with zero $(0, \ldots, 0)$, addition $(b_1, \ldots, b_k) + (b'_1, \ldots, b'_k) = (b_1 + b'_1, \ldots, b_k + b'_k)$, and scalar multiplication $a(b_1, \ldots, b_k) = (a\, b_1, \ldots, a\, b_k)$. In this case we write $B_{i_1} \oplus \ldots \oplus B_{i_k} := \bigoplus_{i \in I} B_i$.

- Let for each $i \in I$,
$$g_i : \prod_{j \in I} B_j \to B_i : g \mapsto g(i). \tag{3.1}$$
Then for any B $\boxed{\mathrm{M}}/A$ and $\{f_i : B \to B_i\}$, f_i $\textcircled{1}/A$ for all $i \in I$, there exists a unique $f : B \to \prod_{i \in I} B_i$ $\textcircled{1}/A$ such that for all $i \in I$ we have that $f_i = g_i \circ f$.

- Let for each $i \in I$,
$$g_i : B_i \to \bigoplus_{j \in I} B_j : b \mapsto \left(j \mapsto \left\{ \begin{array}{ll} b & i = j \\ 0_{B_j} & i \neq j \end{array} \right. \right). \tag{3.2}$$
Then for any B $\boxed{\mathrm{M}}/A$ and $\{f_i : B_i \to B\}$, f_i $\textcircled{1}/A$ for all $i \in I$, there exists a unique $f : \bigoplus_{i \in I} B_i \to B$ $\textcircled{1}/A$ such that for all $i \in I$ we have that $f_i = f \circ g_i$.

Proof.
- This is clear from the definition.

- If I is finite, then the set of $i \in I$ for which $g(i) \neq 0_{B_i}$ is always finite for any $g \in \prod_{i \in I} B_i$. Hence $\prod_{i \in I} B_i = \bigoplus_{i \in I} B_i$. We can identify this set with $B_{i_1} \times \ldots \times B_{i_k}$ using the maps $(b_1, \ldots, b_k) \mapsto (I \to \bigcup_{i \in I} B_i : i \mapsto b_l$ for which $i = i_l$) and $g \mapsto (g(i_1), \ldots, g(i_k))$.

- Suppose $f : B \to \prod_{i \in I} B_i$ is $\textcircled{1}/A$ and satisfies $f_i = g_i \circ f$ for all $i \in I$. Then for $b \in B$ and $i \in I$ we have $f(b)(i) = g_i(f(b)) = f_i(b)$, which uniquely determines $f(b) \in \prod_{i \in I} B_i$. Now define for $b \in B$, $i \in I$ by $f(b)(i) := f_i(b)$, then $f_i(b) = f(b)(i) = g_i(f(b))$, so $f_i = g_i \circ f$. Furthermore, f $\textcircled{1}/A$ as all f_i $\textcircled{1}/A$. Hence f is the desired unique map.

- Let $g \in \bigoplus_{i \in I} B_i$, then there exist finitely many $i_1, \ldots, i_k \in I$ such that $g(i_l) \neq 0_{B_i}$ for $1 \leq l \leq k$. Define $f(g) := f_{i_1}(g(i_1)) + \ldots + f_{i_k}(g(i_k))$ to obtain a map $f : \bigoplus_{i \in I} B_i \to B$. Since all f_i $\textcircled{1}/A$ we see that f $\textcircled{1}/A$ by definition. Furthermore, for $i \in I$, $b \in B_i$ we have by definition of f and g_i that $f(g_i(b)) = f_i(g_i(b)(i)) = f_i(b)$, hence $f_i = f \circ g_i$ as desired. Uniqueness of f is also apparent: for a given $g \in \bigoplus_{i \in I} B_i$ we have for the $i_1, \ldots, i_k \in I$ where $g(i_l) \neq 0$ that $g = g_{i_1}(g(i_1)) + \ldots + g_{i_k}(g(i_k))$, so $f(g) = f(g_{i_1}(g(i_1))) + \ldots + f(g_{i_k}(g(i_k))) = f_{i_1}(g(i_1)) + \ldots + f_{i_k}(g(i_k))$ which uniquely determines $f(g)$.

$\square$

By Lemma (3.3.9) we see that finite direct products or sums cannot be distinguished. We will often abbreviate $B^k := \underbrace{B \oplus \ldots \oplus B}_{k} \simeq \underbrace{B \times \ldots \times B}_{k}$.

Lemma 3.3.10: Factorisation
Let A $\boxed{\mathrm{R}}$, B, C $\boxed{\mathrm{M}}/A$.

Then for any $f : B \to C$ $\textcircled{1}/A$ there exists an unique $g : B/\ker f \to C$ $\textcircled{1}/A$ such that $f(b) = g([b])$ for all $b \in B$. Furthermore, $\ker g = \{[0]\}$ and
$$B/\ker f \simeq f(B) \leq C.$$

Proof. Let $f : B \to C$ $\textcircled{l}$ $/A$ be given. Define $g : B/\ker f \to C$ by $g([b]) := f(b)$ for all $b \in B$. Then this map is well-defined, for if $[b_1] = [b_2]$ then $b_1 - b_2 \in \ker f$, so $g([b_1]) - g([b_2]) = f(b_1) - f(b_2) = f(b_1 - b_2) = 0$, so $g([b_1]) = g([b_2])$. It is also linear because f is linear, and unique by definition.

As g $\textcircled{l}$, $[0] \in \ker g$. Let $[b] \in \ker g$ be arbitrary, then $g([b]) = f(b) = 0$, so $b \in \ker f$, but then $[b] = [0]$, hence $\ker g = \{[0]\}$.

Because g is injective, and $f(b) = g([b])$ from which we see that $g(B/\ker f) = f(B)$, so $B/\ker f \to f(B) : [b] \mapsto g([b])$ is a $\boxed{\mathbf{M}}$-isomorphism. $\square$

Lemma 3.3.11

Let A $\boxed{\mathbf{R}}$ and B, C, D $\boxed{\mathbf{M}}$ $/A$.

The we have the following $\boxed{\mathbf{M}}$-isomorphisms.

- $(B \oplus C) \oplus D \simeq B \oplus (C \oplus D) \simeq B \oplus C \oplus D$,
- $B \oplus C \simeq C \oplus B$,
- $(B \oplus C)/C \simeq B$, if furthermore $C \leq B$, then also $(B/C) \oplus C \simeq B$.

Proof.

- Consider the isomorphisms $((b, c), d) \mapsto (b, (c, d))$ and $(b, (c, d)) \mapsto (b, c, d)$.
- Consider $(b, c) \mapsto (c, b)$.
- Consider $B \oplus C \to B : (b, c) \to b$ (the kernel of this map is precisely $\{0\} \times C \simeq C$ and it is surjective) and apply Lemma (3.3.10) to obtain that $(B \oplus C)/C \simeq B$. On the other hand if $C \leq B$ we can take $B \oplus C \to (B/C) \oplus C : (b, c) \mapsto ([b], c)$ (note that $([b], c) = ([0], 0)$ if and only if $c = 0$ and $b \in C$, so the kernel of this map is $C \times \{0\} \simeq C$) and we then use Lemma (3.3.10) to obtain that $B \simeq (B \oplus C)/C \simeq (B/C) \oplus C$.

$\square$

Definition 3.3.12: Dimension

Let A $\boxed{\mathbf{R}}$, B $\boxed{\mathbf{M}}$ $/A$, $k \in \mathbb{N}$.

We define the *span* of a collection $\{b_1, \ldots, b_k\} \subseteq B$ to be

$$\langle b_1, \ldots, b_k \rangle_A := \{a_1\, b_1 + \ldots + a_k\, b_k \in B | a_1, \ldots, a_k \in A\} \leq B.$$

We call a collection $\{b_1, \ldots, b_k\} \subseteq B$ *linearly independent* (or say that $b_1, \ldots, b_k$ *are linearly independent*) if for all $a_1, \ldots, a_k \in A$ we have that

$$a_1\, b_1 + \ldots + a_k\, b_k = 0 \to a_1 = \ldots = a_k = 0.$$

If $\{b_1, \ldots, b_k\}$ is not linearly independent, it is called *linearly dependent*. An empty collection $\emptyset \subseteq B$ is never linearly independent. We call infinite subset $\{b_i \in B | i \in I\}$ linearly independent if for all $k \in \mathbb{N}$, $i_1, \ldots, i_k \in I$, the collection $\{b_{i_1}, \ldots, b_{i_k}\} \subseteq B$ is linearly independent.

We say that B *has dimension* 0 if $B = \{0\}$.

If there exists a $k \in \mathbb{N}$ and $b_1, \ldots, b_k \in B$ that are linearly independent and for which $B = \langle b_1, \ldots, b_k \rangle_A$, we say that B *has dimension* k.

We say that B *has finite dimension* if there exists a $k \in \mathbb{N}_0$ such that B has dimension k, otherwise we say that B *has infinite dimension*.

Example 3.3.13
Let $k \in \mathbb{N}$, then for

$$e_1 := (1, 0, \ldots, 0), e_2 := (0, 1, \ldots, 0), \ldots, e_k := (0, 0, \ldots, 1) \in \mathbb{K}^k$$

(these are called the *(canonical) basis vectors of* $\mathbb{K}^k$) we have that $\{e_1, \ldots, e_k\}$ is linearly independent.

For let $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, then $0 = (0, \ldots, 0) = \alpha_1 e_1 + \ldots + \alpha_k e_k = (\alpha_1, \ldots, 0) + \ldots + (0, \ldots, \alpha_k) = (\alpha_1, \ldots, \alpha_k)$ if and only if $\alpha_1 = \ldots = \alpha_k = 0$.

Furthermore, by definition $\mathbb{K}^k = \langle e_1, \ldots, e_k \rangle_{\mathbb{K}}$, so $\mathbb{K}^k$ has dimension k.

Lemma 3.3.14
Let A **R**, B **M**$/A$.

If $\langle b_1, \ldots, b_k \rangle_A \subseteq \langle b'_1, \ldots, b'_l \rangle_A$ for $b_1, \ldots, b_k$ and $b'_1, \ldots, b'_l$ both linearly independent, then $k \le l$.

In particular if B has dimension k and B has dimension l, then $k = l$, and any collection of linearly independent vectors of B has at most k elements.

Furthermore, if C **M**$/A$ and $B \simeq C$, then B has dimension k if and only if C has dimension k.

Proof. Suppose that for any $1 \le m \le l$ we have that $\{b'_m, b_2, \ldots, b_k\}$ is linearly dependent. Then for each m, $b'_m \in \langle b_2, \ldots, b_k \rangle_A$. However, $b_1 \in \langle b'_1, \ldots, b'_l \rangle_A$, so this would mean that $b_1 \in \langle b_2, \ldots, b_k \rangle_A$ and therefore $\{b_1, \ldots, b_k\}$ is linearly dependent, which leads to a contradiction.

Hence there is some $m_1 \in \{1, \ldots, l\}$ such that $\{b'_{m_1}, b_2, \ldots, b_k\}$ is linearly independent. Following the same reasoning for $b_2, b_3, \ldots, b_k$ we find $m_2, m_3, \ldots, m_k \in \{1, \ldots, l\}$ such that $\{b'_{m_1}, b'_{m_2}, b'_{m_3}, \ldots, b'_{m_k}\}$ is linearly independent. Hence all $m_1, \ldots, m_k$ must be distinct and therefore $k \le l$.

This makes the dimension of B unique: if B has dimension k and B has dimension l then by the above $k \le l$ and $l \le k$, so $k = l$.

Let C **M**$/A$, $B \simeq C$ and suppose B has dimension k. Let $f : B \to C$ be the **M**-isomorphism and let $b_1, \ldots, b_k \in B$ be linearly independent such that $B = \langle b_1, \ldots, b_k \rangle_A$. Now let $c_l := f(b_l)$ for $1 \le l \le k$. Then $c_1, \ldots, c_k \in C$ are linearly independent: suppose $a_1\, c_1 + \ldots + a_k\, c_k = 0$ for $a_1, \ldots, a_k \in A$. Then (f **l**$/A$) $f(a_1\, b_1 + \ldots + a_k\, b_k) = 0$, but f is injective and therefore $\ker f = \{0\}$, hence $a_1\, b_1 + \ldots + a_k\, b_k = 0$ and (linear independence) therefore $a_1 = \ldots = a_k = 0$. Furthermore, any $c \in C = f(B)$ can be written as $c = f(b)$ for some $b \in B$, but $b = \langle b_1, \ldots, b_k \rangle_A$, so there exist $a_1, \ldots, a_k \in A$ such that $b = a_1\, b_1 + \ldots + a_k\, b_k$ and hence $c = f(b) = a_1\, f(b_1) + \ldots + a_k\, f(b_k) = a_1\, c_1 + \ldots + a_k\, c_k \in \langle c_1, \ldots, c_k \rangle_A$. Therefore $C = \langle c_1, \ldots, c_k \rangle_A$ and hence C has dimension k. For the converse, simply exchange B and C. □

Lemma 3.3.15
Let A **R**.

If B **M**$/A$ has dimension k for some $k \in \mathbb{N}$, then $B \simeq A^k$ as **M**$/A$.

In particular, if B, C **M**$/A$, B has dimension k, and C has dimension l, then the dimension of $B \oplus C$ equals $k + l$.

Proof. Suppose B has dimension k, then there exists $b_1, \ldots, b_k \in B$ linearly independent such that $B = \langle b_1, \ldots, b_k \rangle_A$. Now consider the map $f : A^k \to B$ given by $f(a_1, \ldots, a_k) := a_1\, b_1 + \ldots + a_k\, b_k$. Then this map is surjective because $B = \langle b_1, \ldots, b_k \rangle_A = f(A)$. Suppose $f(a_1, \ldots, a_k) = f(a'_1, \ldots, a'_k)$, then $0 =$

$a_1\, b_1 + \ldots + a_k\, b_k - a'_1\, b_1 - \ldots - a'_k\, b_k = (a_1 - a'_1)\, b_1 + \ldots + (a_k - a'_k)\, b_k$. As $b_1, \ldots, b_k$ are linearly independent we therefore have that $a_1 - a'_1 = \ldots = a_k - a'_k = 0$, so $(a_1, \ldots, a_k) = (a'_1, \ldots, a'_k)$. Hence f is injective.

Therefore f is bijective, and by definition f **l** $/A$, so f is a **M**-isomorphism and hence $B \simeq A^k$.

Suppose B has dimension k and C **M** $/A$ has dimension l, then $B \simeq A^k$, $C \simeq A^l$ and therefore (Lemma (3.3.11)) we have $B \oplus C \simeq (A^k) \oplus (A^l) \simeq A^{k+l}$ and therefore by Lemma (3.3.14) the dimension of $B \oplus C$ equals the dimension of A^{k+l} which is $k + l$. □

Lemma 3.3.16: Rank lemma

Let A **R**, B, C **M** $/A$, and $f : B \to C$ **l** $/A$.

If B has dimension k, then the dimension of $\ker f$ plus the dimension of $f(B)$ equals k.

Proof. Using Lemma (3.3.10) and Lemma (3.3.11) we find $B \simeq (B/\ker f) \oplus \ker f \simeq f(B) \oplus \ker f$. Hence we can consider $f(B) \leq B$, which must have finite dimension as B has finite dimension (Lemma (3.3.14)), now the result follows from Lemma (3.3.15). □

Definition 3.3.17: Algebraic dual of a module

Let A **R** and B **M** $/A$.

Then the *(algebraic) dual of* B is defined to be the set

$$B^* := \{f : B \to A | f \text{ (l) } /A\},$$

together with $0 : b \mapsto 0_A$, $(f+g)(b) := f(b) + g(b)$, and $(a\, f)(b) := a\, f(b)$, which make B^* **M** $/A$.

Definition 3.3.18: Algebraic dual

Let A **R**, B, C **M** $/A$.

Then for any $f : B \to C$ **l** $/A$ we define its algebraic dual, or *transpose* (which is again **l** $/A$) as

$$f^* : C^* \to B^* : g \mapsto g \circ f.$$

We can directly see that if B has dimension k, then so does B^*: let $b_1, \ldots, b_k \in B$ be linearly independent, then the maps $B \to A$ defined for each $1 \leq l \leq k$ by mapping any $b = a_1\, b_1 + \ldots + a_k\, b_k$ to a_l are linearly independent in B^* (simply apply them to $b_1, \ldots, b_k$).

Theorem 3.3.19: Finite dimensional duality

Let A **R**, B **M** $/A$.

Denote for any $C \leq B$ the set

$$C^\perp := \{f \in B^* | \forall c \in C : f(c) = 0\} \leq B^*,$$

and for any $D \leq B^*$ the set

$$D^\perp := \{b \in B | \forall f \in D : f(b) = 0\} \leq B.$$

- Then for all $C \leq B$, $D \leq B^*$ we have

$$C \leq (C^\perp)^\perp, \qquad D \leq (D^\perp)^\perp.$$

- For all $C_1 \leq C_2 \leq B$, $D_1 \leq D_2 \leq B^*$ we have
$$C_1^\perp \geq C_2^\perp, \qquad D_1^\perp \geq D_2^\perp.$$
- For all $C \leq B$, $C^\perp$ has dimension k if and only if B/C has dimension k.
- For all $D \leq B^*$, D has dimension k if and only if $B/D^\perp$ has dimension k.
- The correspondences $C \mapsto C^\perp$, $D \mapsto D^\perp$ are each other's inverse and form a bijection between the collection of all $C \leq B$ for which B/C has finite dimension and the collection of all $D \leq B^*$ of finite dimension.

  In particular, if C and D have finite dimension:
$$B/C = B/(C^\perp)^\perp, \qquad D = (D^\perp)^\perp.$$

Proof. We follow [Bou1947] (Chapitre II, par. 4, no. 6, Théorème 1).

- Let $c \in C$, $f \in C^\perp$, then (definition of $C^\perp$) $f(c) = 0$, hence $\forall f \in C^\perp$: $f(c) = 0$ and therefore $c \in (C^\perp)^\perp$, so $C \leq (C^\perp)^\perp$. Similarly $D \leq (D^\perp)^\perp$.
- Suppose $C_1 \leq C_2$, let $f \in C_2^\perp$, then $f(c) = 0$ for all $c \in C_2 \geq C_1$, so in particular $f(c) = 0$ for all $c \in C_1$, hence $f \in C_1^\perp$. This means $C_1^\perp \geq C_2^\perp$ and similarly $D_1^\perp \geq D_2^\perp$.
- First of all note that $(B/C)^* \simeq C^\perp$ via $(B/C)^* \to C^\perp : f \mapsto (B \to A : b \mapsto f([b]))$ and its inverse (well-defined because of Lemma (3.3.10)) $C^\perp \to (B/C)^* : g \mapsto (B/C \to A : [b] \mapsto g(b))$. Therefore, by Lemma (3.3.14), if B/C has dimension k, $(B/C)^*$ has dimension k and hence $C^\perp$ does too.

  Conversely if $C^\perp$ has dimension k, if B/C has finite dimension, then this dimension (by the converse) must be equal to k. Suppose B/C has infinite dimension. Then there exists a $C_1 \leq B$ with $C_1 \geq C$ and B/C_1 of dimension $k+1$. However, then $C_1^\perp \leq C^\perp$ also has dimension $k+1$, which is impossible by Lemma (3.3.14) as C has dimension k. Therefore B/C must have finite dimension.
- Suppose that D has dimension k, then let $f_1, \ldots, f_k \in D$ be linearly independent. Now define $f : B \to A^k : b \mapsto (f_1(b), \ldots, f_k(b))$, then the dimension of $f(B)$ is at most k. On the other hand, $\ker f = D^\perp$, so the dimension of $B/D^\perp$ equals the dimension of $f(B)$ by Lemma (3.3.10) and is therefore at most k. By the previous part, the dimension of $B/D^\perp$ equals the dimension of $(D^\perp)^\perp \geq D$, which is at least k. Hence $B/D^\perp$ has dimension k.

  Conversely if $B/D^\perp$ has dimension k, then $(D^\perp)^\perp \geq D$ has dimension k, so the dimension of D is at most k. Hence D is finite dimensional and therefore (converse) must have dimension equal to that of $B/D^\perp$.
- Suppose $D \leq B^*$ has dimension k, then $B/D^\perp$ has dimension k and hence $(D^\perp)^\perp$ as well. As $D \leq (D^\perp)^\perp$ also has dimension k we therefore have $D = (D^\perp)^\perp$. Suppose that for $C \leq B$, B/C has dimension k, then $C^\perp$ has dimension k and therefore $B/(C^\perp)^\perp$ is k dimensional. Hence $B/C = B/(C^\perp)^\perp$. This makes $C \mapsto C^\perp$ and $D \mapsto D^\perp$ inverse operations.

□

Lemma 3.3.20: Dual of the dual
Let A $\boxed{\text{R}}$ with $0 \neq 1$, B $\boxed{\text{M}}$ $/A$.

Then the map

$$f : B \to (B^*)^* : b \mapsto (g \mapsto g(b))$$

is $\boxed{1}$ $/A$ and injective.

Proof. Let $g, h \in B^*$, $a \in A$, then for any $b \in B$, $f(b)(g + a\,h) = (g + a\,h)(b) = g(b) + a\,h(b) = f(b)(g) + a\,f(b)(h)$, so $f(b) : B^* \to A$ $\boxed{1}$ $/A$ and hence $f(b) \in (B^*)^*$. Let $b_1, b_2 \in B$, $a \in A$, then for any $g \in B^*$, $f(b_1 + a\,b_2)(g) = g(b_1 + a\,b_2) = g(b_1) + a\,g(b_2) = f(b_1)(g) + a\,f(b_2)(g)$, so f $\boxed{1}$ $/A$.

Fix some $b \in B$, $b \neq 0$. Then the space $C := \{a\,b \in B | a \in A\} \leq B$ is $\boxed{\text{M}}$ $/A$. We can define a map $g_C : C \to A$ by $g_C(a\,b) := a$, then $g_C \in C^*$ and $g_C(b) = g_C(1b) = 1 \neq 0$, so $g_C \neq 0 \in C^*$. Create the collection

$$\mathcal{A} := \{g_D : D \to A \mid C \leq D \leq B, g_D \in D^*, g_D|_C = g_C\}$$

partially ordered by $g_D \leq g_E$ if and only if $D \leq E$ and $g_E|_D = g_D$. Then $\mathcal{A} \neq \emptyset$ (since $g_C \in \mathcal{A}$) and satisfies the chain condition (let $\mathcal{A}' \subseteq \mathcal{A}$ be a totally ordered subset, then $\bigcup \mathcal{A}'$ is an upper bound, because all maps in $\mathcal{A}'$ are compatible due to the imposed ordering). By Zorn's lemma we therefore find that $\mathcal{A}$ has a maximal element $g_D \in \mathcal{A}$ with respect to the partial ordering. Suppose $D \neq B$, then there exists a $b_1 \in B \setminus D$, but then we can construct $E := \{d + a\,b_1 \in B | d \in D, a \in A\}$ satisfying $C \leq D \leq E \leq B$ and a map $g_E : E \to A$ by $g_E(d + a\,b_1) = g_D(d) + a$, clearly satisfying $g_E|_D = g_D$ and $g_E \in E^*$, which contradicts maximality of g_D. Therefore necessarily $B = D$ and we find a $g \in B^*$ such that $g(b) = g_C(b) = 1$. So for any $b \in B$, $b \neq 0$ there exists a $g \in B^*$ with $g(b) = 1$ (a miniature version of Theorem (3.3.22)).

In particular, $f(b)(g) = g(b) = 1 \neq 0$, so $f(b) \neq 0$ and hence $b \notin \ker f$ if $b \neq 0$. On the other hand $f(0)(g) = g(0) = 0$ for all $g \in B^*$, so $\ker f = \{0\}$ and hence f is injective. □

Definition 3.3.21: Vector space
Let A $\boxed{\text{R}}$, B $\boxed{\text{M}}$ $/A$.

Then we call B an *A-vector space* (denoted by B $\boxed{\text{Vs}}$ $/A$) if A $\boxed{\text{F}}$.

Theorem 3.3.22: Hahn-Banach ($\mathbb{R}$)
Let A $\boxed{\text{Vs}}$ $/\mathbb{R}$, and $f : A \to \mathbb{R}$, such that for all $a_1, a_2 \in A$, $f(a_1 + a_2) \leq f(a_1) + f(a_2)$ and $f(\alpha\,a_1) = \alpha\,f(a_1)$ for all $\alpha \in [0, \infty[$.

Then for any $g \in B^*$ with $B \leq A$, satisfying $g(b) \leq f(b)$ for all $b \in B$, there exists an $h \in A^*$ such that $h|_B = g$, and $h(a) \leq f(a)$ for all $a \in A$.

Proof. Consider the collection

$$\mathcal{A} := \{h_C : C \to \mathbb{R} \boxed{1} \mid B \leq C \leq A, h_C|_B = g, \forall c \in C : h_C(c) \leq f(c)\},$$

together with the partial ordering $h_C \leq h_D$ if and only if $C \leq D$ and $h_D|_C = h_C$. Then $\mathcal{A}$ satisfies the chain condition (let $\mathcal{A}' \subseteq \mathcal{A}$ be a totally ordered subset, then $\bigcup \mathcal{A}'$ is an upper bound because all maps in $\mathcal{A}'$ are compatible under restrictions since $\mathcal{A}'$ is totally ordered) and $\mathcal{A} \neq \emptyset$ because $g \in \mathcal{A}$, and hence by Zorn's lemma $\mathcal{A}$ has a maximal element with respect to its partial ordering, denote this element by h_C. Suppose $C \neq A$, then there exists an $a \in A \setminus C$ and since $C \leq A$, this implies that for

$$D := \{c + \alpha\, a \mid c \in C, \alpha \in \mathbb{R}\}$$

we have $B \leq C \leq D \leq A$.

Now let $h_D \in D^*$ be any function satisfying $h_D|_C = h_C$ and $h_D(d) \leq f(d)$ for all $d \in D$. Then for any $d = c + \alpha\, a \in D$ we have $h_D(d) = h_D(c + \alpha\, a) = h_D(c) + \alpha\, h_D(a) = h_C(c) + \alpha\, h_D(a)$. Furthermore, if $\alpha = 0$, $h_D(d) = h_C(c) + 0 \leq f(c) = f(d)$ directly (as $h_C \in \mathcal{A}$).

If $\alpha < 0$, $h_D(d) = h_C(c) - |\alpha|\, h_D(a) \leq f(d) = f(c - |\alpha|\, a)$, so $h_C(\frac{1}{|\alpha|}\, c) - h_D(a) \leq f(\frac{1}{|\alpha|}\, c - a)$, so $h_D(a) \geq h_C(c') - f(c' - a)$ for $c' = \frac{1}{|\alpha|}\, c \in C$.

If $\alpha > 0$ we find in a similar way that $h_D(d) = h_C(c) + |\alpha|\, h_D(a) \leq f(c + |\alpha|\, a)$, so $h_D(a) \leq f(c' + a) - h_C(c')$ for $c' = \frac{1}{|\alpha|}\, c \in C$.

So if $h_D \in D^*$ satisfies $h_D|_C = h_c$ and $h_D(d) \leq f(d)$ for all $d \in D$, we have that necessarily $h_D(c + \alpha\, a) = h_C(c) + \alpha\, h_D(a)$ with for all $c, c' \in C$

$$h_C(c) - f(c - a) \leq h_D(a) \leq f(c' + a) - h_C(c') \tag{3.3}$$

since the correspondence $c \leftrightarrow \frac{1}{|\alpha|}\, c$ in C is bijective for $\alpha \neq 0$.

Conversely, for a certain $\beta \in \mathbb{R}$ which satisfies Equation (3.3) we can define for all $d = c + \alpha\, a \in D$ the function $h_D : D \to \mathbb{R}$ by $h_D(c + \alpha\, a) := h_C(c) + \alpha\, \beta$. Then $h_D(a) = \beta$ and $h_D \in D^*$ as $h_C \in C^*$. Furthermore $h_D|_C = h_C$ (case $\alpha = 0$) and $h_D(d) \leq f(d)$ for all $d \in D$ (reverse the reasoning leading to Equation (3.3)).

Now for any $c, c' \in C$ we have $h_C(c) + h_C(c') = h_C(c + c') \leq f(c + c') = f(c - a + c' + a) \leq f(c - a) + f(c' + a)$, so $h_C(c) - f(c - a) \leq f(c' + a) - h_C(c')$ for all $c, c' \in C$. This means that both $\beta_- := \sup\{h_C(c) - f(c - a) \in \mathbb{R} | c \in C\}$ and $\beta_+ := \inf\{f(c' + a) - h_C(c') \in \mathbb{R} | c' \in C\}$ exist in $\mathbb{R}$, $\beta_- \leq \beta_+$ and that any β in the nonempty interval $[\beta_-, \beta_+]$ satisfies Equation (3.3).

So there exist $\beta_-, \beta_+ \in \mathbb{R}$, $\beta_- \leq \beta_+$ such that for all $\beta \in [\beta_-, \beta_+]$ the function $h_D : D \to \mathbb{R}$ defined by $h_D(c + \alpha\, a) = h_C(c) + \alpha\, \beta$ satisfies $h_D \in D^*$, $h_D|_C = h_C$ and $h_D(d) \leq f(d)$ for all $d \in D$. Therefore $h_D \in \mathcal{A}$ and since $D \geq C$, $D \neq C$ we have $h_D > h_C$ contradicting the maximality of h_C.

Therefore the assumption $C \neq A$ leads to a contradiction: necessarily $C = A$ and hence the maximal element $h_A =: h$ is the sought after extension of g. $\square$

It is tempting to try and prove the Hahn-Banach theorem for other fields besides $\mathbb{R}$. Later (Theorem (4.2.6)) we will see that we can also extend Theorem (3.3.22) to $\mathbb{C}$, but the following example shows that completeness of the considered field is very important.

Example 3.3.23: Hahn-Banach fails over $\mathbb{Q}$

Here we will investigate the particulars of Theorem (3.3.22) where instead of vector spaces over $\mathbb{R}$ we consider vector spaces over the field $\mathbb{Q}$.

Consider $A = \mathbb{R}^2$ considered as Vs $/\mathbb{Q}$ and $B = \mathbb{Q} \leq A$ (identified with $\mathbb{Q} \times \{0\}$ via $x \mapsto (x, 0)$). Choose $f : A \to \mathbb{R} : (x, y) \mapsto |x|$, then $f((x_1, y_1) + (x_2, y_2)) = |x_1 + x_2| \leq |x_1| + |x_2| = f(x_1, y_1) + f(x_2, y_2)$ and for all $\alpha \in \mathbb{Q}$ we have $f(\alpha\, (x_1, y_1)) = |\alpha\, x_1| = |\alpha|\, f(x_1, y_1)$. Pick $g : B \to \mathbb{Q} : x \mapsto x$, then clearly $g \in B^*$ and $g(x) = x \leq |x| = f(x, 0)$ for all $x \in B$.

Let $h \in A^*$ (so $h : \mathbb{R}^2 \to \mathbb{Q}$ is $\textcircled{1}/\mathbb{Q}$) satisfy both $h|_B = g$ and $h(x,y) \leq f(x,y)$ for all $(x,y) \in A$. Consider the linear subspace

$$D := \{(x,0) + \alpha\,(\sqrt{2},1) \in A \mid (x,0) \in B, \alpha \in \mathbb{Q}\} \leq A.$$

Then because h is linear and restricts to g we have $h((x,0) + \alpha\,(\sqrt{2},1)) = h(x,0) + \alpha\,h(\sqrt{2},1) = g(x) + \alpha\,h(\sqrt{2},1) = x + \alpha\,\beta$ for $\beta := h(\sqrt{2},1) \in \mathbb{Q}$. As $h(x,y) \leq f(x,y)$ for all $(x,y) \in A \geq D$ and $B \leq D \leq A$ we see from Equation (3.3) and the proof of Theorem (3.3.22) (with $a = (\sqrt{2},1) \in A \setminus B$) that necessarily for all $x \in \mathbb{Q}$ we have $\beta \leq f((x,0) + (\sqrt{2},1)) - g(x)$ and $\beta \geq g(x) - f((x,0) - (\sqrt{2},1))$. For $x = 2 \in \mathbb{Q}$ we therefore find $\beta \geq g(2) - f((2,0) - (\sqrt{2},1)) = 2 - |2 - \sqrt{2}| = \sqrt{2}$ and for $x = 1 \in \mathbb{Q}$, $\beta \leq f((1,0) + (\sqrt{2},1)) - g(1) = |1 + \sqrt{2}| - 1 = \sqrt{2}$. Therefore $\beta = \sqrt{2}$ and hence $\sqrt{2} = \beta \in \mathbb{Q}$ which leads to a contradiction.

Therefore such a bounded linear extension h of g cannot exist in this case: Theorem (3.3.22) does not hold for vector spaces over $\mathbb{Q}$.

Chapter 4

Topology and algebra

We will now provide the algebraical objects from Chapter 3 with the geometrical properties from Chapter 2, which is necessary to be able to do analysis later in Chapter 5. There we need to investigate exactly how fast the value of a function changes if we 'move around' a fixed point for differentiability, this is not possible in the general setting of Chapter 2. Of particular use in this regard will be the notion of abc subsets in Definition (4.3.4) and Section 4.5.

4.1 Topological modules

We now make a connection between the discussed topological and algebraical concepts.

Definition 4.1.1: Topological group ( )
Let A be a set.

Then we call A a *topological group* (denoted by A) if A such that the multiplication and inversion maps of the group structure are both with respect to the topology on A.

Definition 4.1.2: Topological ring ( )
Let A be a set.

Then we call A a *topological ring* (denoted by A) if A such that the multiplication and addition maps given by the ring structure on A are all with respect to the topology on A.

By Lemma (3.2.3) we know that for topological rings A, the map $A \to A : a \mapsto -a$ is given by the composition of $A \times A \to A : (a_1, a_2) \mapsto a_1 a_2$ with $A \to A \times A : a \mapsto (-1, a)$, which are both , so $a \mapsto -a$. This makes the Abelian group structure on A a .

Definition 4.1.3: Topological module ( )
Let A and B a set.

Then we call B a *topological A-module* (denoted by B $/A$) if B $/A$ such that the scalar multiplication and addition maps given by the A-module structure on B are all with respect to the topology on B.

Definition 4.1.4: Morphisms of topological A-modules
Let A $\mathrm{T_R}$, B, C $\mathrm{T_M}$ $/A$.

Then all maps $f : B \to C$ C I $/A$ are *topological A-module morphisms between B and C* (denoted by f $\mathrm{T_M}$ $/A$-morphism).

The *identity morphism of B* is the map

$$\mathrm{id}_B : B \to B : b \mapsto b.$$

Following the same reasoning as for topological rings we see that for topological modules, the map $b \mapsto -b$ C. Also note that for B $\mathrm{T_M}$ $/A$ and C T we have that $\{f : C \to B | f \text{ C}\}$ $\mathrm{T_M}$ $/A$ because addition and scalar multiplication are C.

Lemma 4.1.5
Let A $\mathrm{T_R}$, B $\mathrm{T_M}$ $/A$.

Let $V \subseteq B$ be open (resp. closed). Then for all $b \in B$,

$$V + b := \{b_1 + b \in B \mid b_1 \in V\}$$

is open (resp. closed) and for all $a \in A^*$,

$$aV := \{a\, b_1 \in B \mid b_1 \in V\}$$

is open (resp. closed).

Proof. Direct from the fact that for fixed $b \in B$ and $a \in A^*$, the maps $B \to B : b_1 \mapsto b_1 + b$, $B \to B : b_1 \mapsto a\, b_1$ (as B is a topological A-module they are both C) are T-isomorphisms, because they have C inverses $b_1 \mapsto b_1 - b$ and $b_1 \mapsto a^{-1}\, b_1$ as $a \in A^*$. □

Note that we can translate any basis of neighbourhoods from $0 \in B$ to any point $b \in B$ by Lemma (4.1.5) and that the entire topology of B is generated by translations of this basis. This allows us to compare topologies of topological modules more easily.

Lemma 4.1.6: Comparing topologies
Let A $\mathrm{T_R}$, B M $/A$. Let $\mathcal{B}_1, \mathcal{B}_2 \subseteq \mathcal{P}(B)$ be topologies on B such that B $\mathrm{T_M}$ $/A$ with each of these topologies. Let $\mathcal{B}_3$ be a basis of open neighbourhoods of 0 in B with respect to the topology $\mathcal{B}_1$.

If for each $V_1 \in \mathcal{B}_3$ there exists a $V_2 \in \mathcal{B}_2$ such that $V_2 \subseteq V_1$, then $\mathcal{B}_1 \subseteq \mathcal{B}_2$.

Furthermore $\mathcal{B}_1 = \mathcal{T}(\{V + b | V \in \mathcal{B}_3, b \in B\})$ (any basis of open neighbourhoods of 0 in B generates the entire topology via translation).

Proof. Suppose that for each $V_1 \in \mathcal{B}_3$ there exists a $V_2 \in \mathcal{B}_2$ such that $V_2 \subseteq V_1$. Let $V \in \mathcal{B}_1$, then for any $b \in V$, $V - b$ is an open neighbourhoods of 0, so there exists some $V_1 \in \mathcal{B}_3$ such that $0 \in V_1 \subseteq V - b$. By assumption there exists a $V_b \in \mathcal{B}_2$ such that $0 \in V_b \subseteq V_1 \subseteq V - b$. But then $V = \bigcup_{b \in V} \{b\} \subseteq \bigcup_{b \in V} (V_b + b) \subseteq V$, so $V = \bigcup_{b \in V} (V_b + b) \in \mathcal{B}_2$, as all $V_b + b \in \mathcal{B}_2$, which is a topology. So $\mathcal{B}_1 \subseteq \mathcal{B}_2$.

Denote $\mathcal{B}_4 = \mathcal{T}(\{V + b | V \in \mathcal{B}_3, b \in B\})$. First of all note (with Lemma (4.1.5)) that the collection of all $V + b$ for $V \in \mathcal{B}_3$ and $b \in B$ is contained in $\mathcal{B}_1$, which is a topology. Hence $\mathcal{B}_4 \subseteq \mathcal{B}_1$. In a similar way as above (write open set as union of translated basis elements) we find $\mathcal{B}_1 \subseteq \mathcal{B}_4$ and hence $\mathcal{B}_1 = \mathcal{B}_4$: the topology of B is generated by translations of any basis of open neighbourhoods of 0 in B. □

Lemma 4.1.7

Let A $\boxed{\mathrm{T_R}}$, B $\boxed{\mathrm{T_M}}$ $/A$.

Then B $\boxed{\mathrm{T1}}$ if and only if B $\boxed{\mathrm{T2}}$.

Proof. Suppose B $\boxed{\mathrm{T2}}$, then by Theorem (2.2.5) B $\boxed{\mathrm{T1}}$.

Suppose B $\boxed{\mathrm{T1}}$. By Lemma (4.1.5) it is sufficient to only consider the points $0 \in B$ and $b \in B$, $b \neq 0$. As B $\boxed{\mathrm{T1}}$, we have by Lemma (2.2.2) that the set $V := B \setminus \{b\} \subseteq B$ is open. Furthermore, as $0 \neq b$, $0 \in V$, so V is an open neighbourhood of 0 in B. Because B $\boxed{\mathrm{T_M}}$, $B \times B \to B : (b_1, b_2) \mapsto b_1 - b_2$ $\boxed{\mathrm{C}}$. As $0 - 0 = 0$ we therefore have for $0 \in V$ open that there exist $V_1, V_2 \subseteq B$ open such that $0 \in V_1$, $0 \in V_2$ and for all $b_1 \in V_1$, $b_2 \in V_2$ we have $b_1 - b_2 \in V$. Now V_1 is an open neighbourhood of 0 in B and by Lemma (4.1.5), $b + V_2$ is an open neighbourhood of b in B.

Suppose $V_1 \cap (b + V_2) \neq \emptyset$, then there exists a $b_1 \in V_1 \cap (b + V_2)$, and as $b_1 \in b + V_2$ there is a $b_2 \in V_2$ such that $b_1 = b + b_2$. However, then $b_1 \in V_1$, $b_2 \in V_2$, so $b = b_1 - b_2 \in V = B \setminus \{b\}$: we have reached a contradiction.

Therefore necessarily $V_1 \cap (b + V_2) = \emptyset$, so 0 and b can be separated by disjoint open sets for any $b \neq 0$, hence B $\boxed{\mathrm{T2}}$. □

Definition 4.1.8: Topological quotient module

Let A $\boxed{\mathrm{T_R}}$, B $\boxed{\mathrm{T_M}}$ $/A$, $C \leq B$.

Then the *topological quotient module* B/C is the quotient module B/C (where B and C are just considered as modules) together with the quotient topology (Definition (2.1.26)) of $B \to B/C : b \mapsto [b]$.

Lemma 4.1.9: Properties of the topological quotient module

Let A $\boxed{\mathrm{T_R}}$, B $\boxed{\mathrm{T_M}}$ $/A$, $C \leq B$.

Then B/C $\boxed{\mathrm{T_M}}$ $/A$ and B/C $\boxed{\mathrm{T2}}$ if and only if $C \subseteq B$ is closed.

Furthermore, the map $B \to B/C : b \mapsto [b]$ is open.

Proof. We already know that B/C $\boxed{\mathrm{M}}$ $/A$ as quotient module without topological structure, so we only need to verify that addition and scalar multiplication are continuous. Denote the projection map as $f : B \to B/C : b \mapsto [b]$.

Let $E \subseteq B$. Let $b \in f^{-1}(f(E))$, then $f(b) \in f(E)$, so there exists an $e \in E$ with $[b] = f(b) = f(e) = [e]$, so $b = e + c$ for some $c \in C$, but then $b = e + c \in E + C$. Hence $f^{-1}(f(E)) \subseteq E + C$. Conversely, let $b \in E + C$, then $b = e + c$ for some $e \in E$ and $c \in C$, therefore $f(b) = f(e+c) = [e+c] = [e] + [c] = [e] + [0] = [e] \in f(E)$, so $b \in f^{-1}(f(E))$. Therefore $E + C \subseteq f^{-1}(f(E))$. Since f is also surjective, we obtain for any $D \subseteq B/C$ and $E \subseteq B$ that

$$f(f^{-1}(D)) = D \qquad f^{-1}(f(E)) = E + C = \{e + c | e \in E, c \in C\}.$$

We are now going to show that addition is $\boxed{\mathrm{C}}$. Let $[b_1], [b_2] \in B/C$ be arbitrary, and W any open neighbourhood of $[b_1] + [b_2]$ in B/C. Then $f(b_1 + b_2) = [b_1 + b_2] = [b_1] + [b_2] \in W$, so $V := f^{-1}(W) \subseteq B$ is an open neighbourhood of $b_1 + b_2$ in B. As $+$ is continuous on B and $B \times B$ has the product topology, there exist $V_1, V_2 \subseteq B$ open such that $b_1 \in V_1$, $b_2 \in V_2$ and for all $b_3 \in V_1$, $b_4 \in V_2$ we have $b_3 + b_4 \in V$. By Lemma (4.1.5) we have that $V_1 + C = \bigcup_{c \in C} V_1 + c \subseteq B$ is open and similarly $V_2 + C$ is open. Now define $W_1 := f(V_1 + C)$, $W_2 := f(V_2 + C)$, then $W_1, W_2 \subseteq B/C$ are open because of the quotient topology and the fact that $f^{-1}(W_1) = f^{-1}(f(V_1 + C)) = (V_1 + C) + C = V_1 + C$, which is open, and $f^{-1}(W_2) = V_2 + C$ is open. Let $[b_3] \in W_1$, $[b_4] \in W_2$, then $f(b_3) \in W_1$, so $b_3 \in f^{-1}(W_1) = V_1 + C$, which means that $b_3 = b_5 + c_5$ for some $b_5 \in V_1$ and $c_5 \in C$. Similarly $b_4 = b_6 + c_6$ for some $b_6 \in V_2$ and $c_6 \in C$. Therefore $b_3 + b_4 = (b_5 + c_5) + (b_6 + c_6) = (b_5 + b_6) + (c_5 + c_6) \in V + C$ since $b_5 \in V_1$ and $b_6 \in V_2$. Now $V + C = f^{-1}(f(V)) = f^{-1}(f(f^{-1}(W))) = f^{-1}(W) = V$, so $b_3 + b_4 \in V$ and hence $[b_3] + [b_4] = f(b_3 + b_4) \in f(V) \subseteq W$. So for all open neighbourhoods W of $[b_1] + [b_2]$ in B/C there exist open neighbourhoods W_1, W_2 of $[b_1]$ resp. $[b_2]$ in B/C such that for all $[b_3] \in W_1$, $[b_4] \in W_2$ we have $[b_3] + [b_4] \in W$. This makes addition on B/C $\boxed{\text{C}}$.

Similarly scalar multiplication $\boxed{\text{C}}$ and therefore B/C $\boxed{\text{TM}}$$/A$.

Note that B/C $\boxed{\text{T1}}$ if and only if (Lemma (2.2.2)) for all $[b] \in B/C$ we have that $\{[b]\} \subseteq B/C$ is closed, which is the case if and only if (Lemma (4.1.5), $\{[b]\} = \{[0]\} + b)$ $\{[0]\} \subseteq B/C$ is closed. By the quotient topology, $W \subseteq B/C$ is open if and only if $f^{-1}(W) \subseteq B$ is open, which implies that $\{[0]\} \subseteq B/C$ is closed if and only if $f^{-1}(\{[0]\}) = C \subseteq B$ is closed. Therefore B/C $\boxed{\text{T1}}$ if and only if $C \subseteq B$ is closed and hence by Lemma (4.1.7) B/C $\boxed{\text{T2}}$ if and only if $C \subseteq B$ is closed.

Now let $U \subseteq B$ be open, then $f^{-1}(f(U)) = U + C = \bigcup_{c \in C}(U + c)$ which is open (Lemma (4.1.5)) as a union of open sets and therefore $f(U) \subseteq B/C$ is open by definition of the quotient topology. Hence f is an open map. $\square$

Definition 4.1.10: Direct product

Let A $\boxed{\text{TR}}$, $\{B_i | i \in I\}$ with B_i $\boxed{\text{TM}}$$/A$ for all $i \in I$.

Then we consider $\prod_{i \in I} B_i$ as a $\boxed{\text{TM}}$$/A$ with the initial topology (Definition (2.1.18)) of the maps $\prod_{j \in I} B_j \to B_i$ defined by Equation (3.1) for all $i \in I$.

Definition 4.1.11: Direct sum

Let A $\boxed{\text{TR}}$, $\{B_i | i \in I\}$ with B_i $\boxed{\text{TM}}$$/A$ for all $i \in I$.

Then we consider $\bigoplus_{i \in I} B_i$ as a $\boxed{\text{TM}}$$/A$ with the final topology (Definition (2.1.23)) of the maps $B_i \to \bigoplus_{j \in I} B_j$ defined by Equation (3.2) for all $i \in I$.

Definition 4.1.12: Topological module dual

Let A $\boxed{\text{TR}}$, B $\boxed{\text{TM}}$$/A$.

Then the *(topological) dual of* B is defined to be

$$B' := \{f \in B^* \mid f \boxed{\text{C}}\},$$

together with the initial topology (Definition (2.1.18)) of $\{B' \to A : f \mapsto f(b) | b \in B\}$. This makes B' $\boxed{\text{TM}}$$/A$.

That B' $\boxed{\text{TM}}$$/A$ follows directly from the fact that $B' \leq B^*$ (from the fact that B^* $\boxed{\text{M}}$$/A$ together with continuity of addition and scalar multiplication in B), which gives B' $\boxed{\text{M}}$$/A$. Furthermore

$$\begin{tikzcd} B' \times B' \arrow[r, "+"] \arrow[d, "(f,g)\mapsto(f(b),g(b))"'] & B' \arrow[d, "f\mapsto f(b)"] \\ A \times A \arrow[r, "+"'] & A \end{tikzcd} \qquad \begin{tikzcd} A \times B' \arrow[r, "\cdot"] \arrow[d, "(a,f)\mapsto(a,f(b))"'] & B' \arrow[d, "f\mapsto f(b)"] \\ A \times A \arrow[r, "\cdot"'] & A \end{tikzcd}$$

show that addition and scalar multiplication on B' are continuous because of the initial topology.

Example 4.1.13: Linearity does not imply continuity, $B' \subsetneq B^*$
Consider $A = \mathbb{R}$,

$$B = \{f : [-1,1] \to \mathbb{R} \mid f \text{ d } [-1,1]\}$$

together with $(f+g)(x) := f(x) + g(x)$, $(\alpha\, f)(x) := \alpha\, f(x)$, $0(x) := 0$, and the norm $\|\cdot\| : B \to \mathbb{R}$,

$$\|f\| := \sup\{|f(x)| \mid x \in [-1,1]\}$$

which make B TM $/\mathbb{R}$.

Consider the map

$$g : B \to \mathbb{R} : f \mapsto f'(0)$$

then $g \in B^*$, while $g \notin B'$. Certainly $g \in B^*$ because differentiation is linear. However, for the sequence

$$x : \mathbb{N} \to B : k \mapsto \left(x \mapsto \frac{\sin(k^2\, x)}{k}\right)$$

we have for all $k \in \mathbb{N}$ that $g(x_k) = k^2 \cos(k^2\, 0)/k = k$. Therefore $\lim_{k\to\infty} g(x_k) = \lim_{k\to\infty} k$ which does not exist in $\mathbb{R}$, while $\lim_{k\to\infty} \|x_k\| = \lim_{k\to\infty} \sup\{|\sin(k^2\, x)|/|k||x \in [0,1]\} \leq \lim_{k\to\infty} 1/k = 0$, so $\lim_{k\to\infty} x_k = 0 \in B$ does exist. Hence g is not C: $g \notin B'$.

4.2 Normed modules

Definition 4.2.1: Normed ring
Let A R.

Then a *seminorm on* A is a map $|\cdot| : A \to \mathbb{R} : a \mapsto |a|$ satisfying for all $a_1, a_2 \in A$ that

- $|a_1| \geq 0$,
- $|a_1\, a_2| \leq |a_1|\, |a_2|$,
- $|a_1 + a_2| \leq |a_1| + |a_2|$,
- $|1| = 1, \quad |0| = 0$.

If in addition $|a_1| = 0 \to a_1 = 0$, we call $|\cdot|$ a *norm on* A.

A ring (resp. field) A together with a (semi)norm is called a *(semi)normed ring (resp. field)*.

Definition 4.2.2: Topology of a normed ring
Let A be a (semi)normed ring.

Then we consider A as a (pseudo)metric space with (pseudo)metric given by

$$d : A \times A \to \mathbb{R} : (a_1, a_2) \mapsto |a_2 - a_1|.$$

Definition 4.2.3: Normed module (‖·‖)
Let A be a (pseudo)normed ring and B M $/A$.

Then a *seminorm on* B is a map $\|\cdot\| : B \to \mathbb{R} : b \mapsto \|b\|$ satisfying for all $b_1, b_2 \in B$ and $a \in A$ that

- $\|b_1\| \geq 0$,
- $\|a\, b_1\| = |a|\, \|b_1\|$,
- $\|b_1 + b_2\| \leq \|b_1\| + \|b_2\|$.

If in addition $\|b_1\| = 0 \to b_1 = 0$ and A is a normed ring, we call $\|\cdot\|$ a *norm on* B.

An A-module B together with a (semi)norm is called a *(semi)normed* A*-module*.

We denote the fact that an A-module B is a normed A-module by B ‖·‖ $/A$.

Definition 4.2.4: Topology of a normed module

Let A be a (semi)normed ring and B M $/A$ a (semi)normed module.

Then we consider B as a (pseudo)metric space with (pseudo)metric given by

$$d : B \times B \to \mathbb{R} : (b_1, b_2) \mapsto \|b_2 - b_1\|.$$

A topological space is called *(semi)normable* if it is T-isomorphic to a (semi)normed module.

Lemma 4.2.5

Let A be a (semi)normed R and B a (semi)normed M $/A$.

Then A $^{\text{T}}$R and B $^{\text{T}}$M $/A$.

Proof. Let $b_1, b_2 \in B$ and $b_3 \in B_B(b_1, \delta_1)$, $b_4 \in B_B(b_2, \delta_2)$. Then $d_B(b_3+b_4, b_1+b_2) = \|(b_3+b_4)-(b_1+b_2)\| = \|(b_3-b_1)+(b_4-b_2)\| \leq \|b_3-b_1\|+\|b_4-b_2\| < \delta_1+\delta_2$ for all $\delta_1, \delta_2 \in]0, \infty[$. From this we obtain continuity of addition by Lemma (2.5.5). Let $b_1 \in B$, $a_1 \in A$, and $b_2 \in B_B(b_1, \delta_1)$, $a_2 \in B_A(a_1, \delta_2)$. Then $d_B(a_1\, b_1, a_2\, b_2) = \|(a_1\, b_1) - (a_2\, b_2)\| = \|(a_1\, b_1) - (a_2\, b_1) + (a_2\, b_1) - (a_2\, b_2)\| \leq \|(a_1 - a_2)\, b_1\| + \|a_2\, (b_1 - b_2)\| < \delta_2\, \|b_1\| + \delta_1\, |a_2|$ which shows continuity of scalar multiplication.

The proof for A is the same. □

Theorem 4.2.6: Hahn-Banach

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$, A Vs $/\mathbb{K}$, and $\|\cdot\| : A \to \mathbb{R}$ a seminorm.

Then for any $f \in B^*$ with $B \leq A$ satisfying $|f(b)| \leq \|b\|$ for all $b \in B$, there exists a $g \in A^*$ such that $g|_B = f$, and $|g(a)| \leq \|a\|$ for all $a \in A$.

Proof. Suppose $\mathbb{K} = \mathbb{R}$. Then by Theorem (3.3.22) ($f(b) \leq |f(b)| \leq \|b\|$ for all $b \in B$ and $\|\cdot\|$ is a seminorm) there exists a $g \in A^*$ with $g|_B = f$ and $g(a) \leq \|a\|$ for all $a \in A$. Now because g 1, $-g(a) = g(-a) \leq \|-a\| = \|a\|$, so $\pm g(a) \leq \|a\|$ which implies that $|g(a)| \leq \|a\|$ for all $a \in A$. g is the desired function.

Suppose $\mathbb{K} = \mathbb{C}$. Then by Theorem (3.3.22) ($\operatorname{Re} f(b) \leq |f(b)| \leq \|b\|$ for all $b \in B$, regard A and B as Vs $/\mathbb{R}$, possible since $\mathbb{R} \leq \mathbb{C}$) there exists a $g : A \to \mathbb{R}$ 1 $/\mathbb{R}$ with $g|_B = \operatorname{Re} f$ and $g(a) \leq \|a\|$ for all $a \in A$. Choose $h : A \to \mathbb{C}$ by $h(a) := g(a) - i\, g(i\, a)$. Then as g 1 $/\mathbb{R}$ and $h(i\, a) = g(i\, a) - i\, g(-a) = i\,(g(a) - i\, g(i\, a)) = i\, h(a)$ we find that h 1 $/\mathbb{C}$. Therefore $h \in A^*$. Now $\operatorname{Re} f(b) = g(b) = \operatorname{Re} h(b)$ and $\operatorname{Im} f(b) = -\operatorname{Re}(i\, f(b)) = \operatorname{Re} f(-i\, b) = g(-i\, b) = -\operatorname{Re}(i\, h(b)) = \operatorname{Im} h(b)$ for all $b \in B$, so $h|_B = f$. Furthermore $\operatorname{Re} h(a) = g(a) \leq \|a\|$ for all $a \in A$. So for any $a \in A$ with $h(a) \neq 0$, $|h(a)| = \operatorname{Re} |h(a)| = \operatorname{Re}\left(\frac{|h(a)|}{h(a)}\, h(a)\right) = \operatorname{Re} h\left(\frac{|h(a)|}{h(a)}\, a\right) \leq \left\|\frac{|h(a)|}{h(a)}\, a\right\| = \left|\frac{|h(a)|}{h(a)}\right| \|a\| = \|a\|$. And for $a \in A$ with $h(a) = 0$, $|h(a)| = 0 \leq \|a\|$ since $\|a\| \geq 0$ for all $a \in A$. Therefore $|h(a)| \leq \|a\|$ for all $a \in A$. h is the desired function. □

4.3 Topological vector spaces

Definition 4.3.1: Topological field ($\boxed{\mathrm{T_F}}$)
Let A be a set.

Then we call A a *topological field* (denoted by A $\boxed{\mathrm{T_F}}$) if A $\boxed{\mathrm{T}}$ $\boxed{\mathrm{F}}$ such that A $\boxed{\mathrm{T_R}}$ and in addition $A^* \to A^* : a_1 \mapsto a_1^{-1}$ $\boxed{\mathrm{C}}$.

Example 4.3.2
Both $\mathbb{R}$ and $\mathbb{C}$ with their usual topologies are normed topological fields.

Definition 4.3.3: Topological vector space ($\boxed{\mathrm{T_{Vs}}}$)
Let A $\boxed{\mathrm{T_F}}$ and B a set.

Then we call B a *topological A-vector space* (denoted by B $\boxed{\mathrm{T_{Vs}}}$ $/A$) if B $\boxed{\mathrm{T}}$ $\boxed{\mathrm{Vs}}$ $/A$ such that B $\boxed{\mathrm{T_M}}$ $/A$.

The morphisms of topological vector spaces are the same as for topological modules (Definition (4.1.4)).

From now on, we will assume $\mathbb{K}$ $\boxed{\mathrm{T_F}}$ to be either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, and denote $\mathbb{K}$'s elements, called *scalars* because of their role in scalar multiplication, by $\alpha, \beta, \ldots$.

Definition 4.3.4: Abc subsets
Let A $\boxed{\mathrm{T_{Vs}}}$ $/\mathbb{K}$.

Then we call any subset $U \subseteq A$ an *abc subset of A* if U is

absorbent: $A = \bigcup_{\alpha \in]0,\infty[} \alpha\, U$,

balanced: $\forall a_1 \in U : \forall \alpha \in \mathbb{K} : (|\alpha| \leq 1 \to \alpha\, a_1 \in U)$,

and convex: $\forall a_1, a_2 \in U : \forall \alpha \in [0,1] : (\alpha\, a_1 + (1-\alpha)\, a_2 \in U)$.

Lemma 4.3.5: Operations preserving abc
Let A, B $\boxed{\mathrm{T_{Vs}}}$ $/\mathbb{K}$.

The notation a/b/c is meant to indicate that the statements hold for each property (absorbent, balanced, convex) separately.

- For any a/b/c subset C of A and $\alpha \in \mathbb{K}$, $\alpha \neq 0$, we have that $\alpha\, C$ is an a/b/c subset of A.
- For any balanced subset C of A and $\alpha, \beta \in \mathbb{K}$, if $|\alpha| \leq |\beta|$ then $\alpha\, C \subseteq \beta\, C$.
- For any balanced subset C of A, $\mathrm{int}(C) = \bigcup_{\alpha \in \overline{B_{\mathbb{K}}(0,1)} \setminus \{0\}} \alpha\, \mathrm{int}(C)$.
- For any a/b/c subset C of A, $\overline{C}$ is an a/b/c subset of A.
- For any a/b/c subset C of A and $f : A \to B$ $\boxed{1}$ $/\mathbb{K}$ surjective, $f(C)$ is an a/b/c subset of B.
- For any a/b/c subset D of B and $f : A \to B$ $\boxed{1}$ $/\mathbb{K}$, $f^{-1}(D)$ is an a/b/c subset of A.

Proof.
- Let $C \subseteq A$ and $\alpha \in \mathbb{K}$, $\alpha \neq 0$. Suppose C is absorbent. Let $a \in A$ be arbitrary, then there exists a $\beta \in \mathbb{K}$ such that $a \in \beta\, C$, so $a \in \frac{\beta}{\alpha}(\alpha\, C)$, therefore $\alpha\, C$ is absorbent.

Suppose C is balanced. Let $a \in \alpha C$ and $\beta \in \mathbb{K}$, $|\beta| \leq 1$, then $a = \alpha\, a_1$ for some $a_1 \in C$ and hence $\beta\, a = \alpha(\beta\, a_1) \in \alpha C$ as C is balanced, so $\alpha\, C$ is balanced.

Suppose C is convex. Let $a_1, a_2 \in \alpha C$, $\beta \in [0,1]$, then $a_1 = \alpha\, a_3$, $a_2 = \alpha\, a_4$ for $a_3, a_4 \in C$, so $\beta\, a_1 + (1-\beta)\, a_2 = \alpha\,(\beta\, a_3 + (1-\beta)\, a_4) \in \alpha C$ as C is convex, so $\alpha\, C$ is convex.

- Let $C \subseteq A$ balanced, $\alpha, \beta \in \mathbb{K}$, $|\alpha| \leq |\beta|$. The case where $|\alpha| = 0$ is clear: $0\, C = \{0\} \subseteq \beta\, C$, so suppose $|\alpha| > 0$. Let $a \in \alpha\, C$, then there exists an $a_1 \in C$ such that $a = \alpha\, a_1$. Hence $a = \beta \frac{\alpha}{\beta}\, a_1 \in \beta C$ since $|\frac{\alpha}{\beta}| = \frac{|\alpha|}{|\beta|} \leq 1$ and C is balanced. So $\alpha\, C \subseteq \beta\, C$.

- Let $C \subseteq A$ be balanced. Let $\alpha \in \mathbb{K}$, $|\alpha| \leq 1$, $\alpha \neq 0$.

  As C is balanced, $\alpha\, C \subseteq 1\, C = C$, so $\operatorname{int}(\alpha\, C) \subseteq \operatorname{int}(C)$. Since $\alpha \operatorname{int}(C) \subseteq \alpha\, C$ is open by Lemma (4.1.5) we therefore have $\alpha \operatorname{int}(C) \subseteq \operatorname{int}(\alpha\, C) \subseteq \operatorname{int}(C)$. So $\bigcup_{\alpha \in \overline{B_{\mathbb{K}}(0,1)} \setminus \{0\}} \alpha \operatorname{int}(C) \subseteq \operatorname{int}(C)$.

  On the other hand, for $\alpha = 1 \in \overline{B_{\mathbb{K}}(0,1)} \setminus \{0\}$ we have $\alpha \operatorname{int}(C) = \operatorname{int}(C)$.

- Let $C \subseteq A$. Suppose C is absorbent, then as $C \subseteq \overline{C}$, $\overline{C}$ is absorbent.

  In the following use that scalar multiplication and addition c.

  Suppose C is balanced. Let $a \in \overline{C}$, $\alpha \in \mathbb{K}$, $|\alpha| \leq 1$. Note that for all $a_1 \in C$, $\alpha\, a_1 \in C$, so by Lemma (2.1.16) and the fact that $a \in \overline{C}$ we find that $\alpha\, a_1 = \lim_{a_1 \to a} \alpha\, a_1 \in \overline{C}$. Hence $\overline{C}$ is balanced.

  Suppose C is convex. Let $a_1, a_2 \in \overline{C}$, $\alpha \in [0,1]$. For all $a_3, a_4 \in C$ we have $\alpha\, a_3 + (1-\alpha)\, a_4 \in C$, so by Lemma (2.1.16) we find $\alpha\, a_1 + (1-\alpha)\, a_2 = \lim_{(a_3,a_4)\to(a_1,a_2)}(\alpha\, a_3 + (1-\alpha)\, a_4) \in \overline{C}$, so $\overline{C}$ is convex.

- Let $C \subseteq A$, $f : A \to B$ 1 surjective. Suppose C is absorbent, then from surjectivity and linearity, $B = f(A) = f(\bigcup_{\alpha \in]0,\infty[} \alpha\, C) = \bigcup_{\alpha \in]0,\infty[} f(\alpha\, C) = \bigcup_{\alpha \in]0,\infty[} \alpha\, f(C)$, so $f(C)$ is absorbent.

  Suppose C is balanced. Let $b \in f(C)$, $\alpha \in \mathbb{K}$, $|\alpha| \leq 1$. Then there exists an $a \in C$ such that $f(a) = b$ and hence $\alpha\, b = \alpha\, f(a) = f(\alpha\, a) \in f(C)$ as $\alpha\, a \in C$, so $f(C)$ is balanced.

  Suppose C is convex. Let $b_1, b_2 \in f(C)$, $\alpha \in [0,1]$, then there exist $a_1, a_2 \in C$ such that $f(a_1) = b_1$, $f(a_2) = b_2$, so $\alpha\, b_1 + (1-\alpha)\, b_2 = \alpha\, f(a_1) + (1-\alpha)\, f(a_2) = f(\alpha\, a_1 + (1-\alpha)\, a_2) \in f(C)$ as $\alpha\, a_1 + (1-\alpha)\, a_2 \in C$, so $f(C)$ is convex.

- Let $D \subseteq B$, $f : A \to B$ 1. Suppose D is absorbent. Let $a \in A$, then $f(a) \in B$, so there exists an $\alpha \in \mathbb{K}$ such that $f(a) \in \alpha\, D$, so either $\alpha = 0$ and $f(a) = 0$, which implies that $a \in f^{-1}(\{0\}) \subseteq f^{-1}(D)$, or $\alpha \neq 0$ and $f(\frac{1}{\alpha}\, a) \in D$, so $\frac{1}{\alpha}\, a \in f^{-1}(D)$ which gives $a \in \alpha\, f^{-1}(D)$. So $f^{-1}(D)$ is absorbent.

  Suppose D is balanced. Let $a \in f^{-1}(D)$, $\alpha \in \mathbb{K}$, $|\alpha| \leq 1$. Then $f(\alpha\, a) = \alpha\, f(a) \in D$, so $\alpha\, a \in f^{-1}(D)$ and $f^{-1}(D)$ is balanced.

  Suppose D is convex. Let $a_1, a_2 \in f^{-1}(D)$, $\alpha \in [0,1]$. Then $f(\alpha\, a_1 + (1-\alpha)\, a_2) = \alpha\, f(a_1) + (1-\alpha)\, f(a_2) \in D$, so $\alpha\, a_1 + (1-\alpha)\, a_2 \in f^{-1}(D)$ and $f^{-1}(D)$ is convex.

□

Lemma 4.3.6: Absorbent and balanced topological basis
Let A TVS $/\mathbb{K}$.

Then

- there exists a basis of open neighbourhoods $\mathcal{A}$ of 0 in A such that all $U \in \mathcal{A}$ are absorbent and balanced,
- for any neighbourhood U of 0 in A, U is absorbent and there exists a $U_1 \in \mathcal{A}$ such that $U_1 + U_1 \in U$.

Proof. Let U be any open neighbourhood of 0 in A. Let $a \in A$. As $\lim_{k\to\infty} \frac{1}{k}\, a = 0\, a = 0$ (continuity of scalar multiplication), there exists a $k \in \mathbb{N}$ such that $\frac{1}{k}\, a \in U$, hence $a \in k\, U$. Therefore $A = \bigcup_{k\in\mathbb{N}} k\, U \subseteq \bigcup_{\alpha\in]0,\infty[} \alpha\, U \subseteq A$, so U is absorbent.

Note that this makes any neighbourhood of 0 in A absorbent.

Since $\lim_{(\alpha,a)\to(0,0)} \alpha\, a = 0$, there exists a $\delta \in]0,\infty[$ and an open neighbourhood U_1 of 0 in A such that for all $\alpha \in B_{\mathbb{K}}(0,\delta)$, and $a \in U_1$ we have $\alpha\, a \in U$. In particular for all $\alpha \in \overline{B_{\mathbb{K}}(0,1)}$, $a \in U_1$, $\alpha\,(\frac{\delta}{2}\, a) \in U$. Hence $0 \in U_2 := \bigcup_{\alpha\in\overline{B_{\mathbb{K}}(0,1)}} \frac{\alpha\delta}{2}\, U_1 \subseteq U$ and by Lemma (4.1.5), U_2 is open as a union of open sets. Furthermore, U_2 is balanced by definition.

So for any neighbourhood U of 0 in A there exists a balanced open neighbourhood U_2 of 0 in A such that $0 \in U_2 \subseteq U$.

Hence we can construct a basis of open neighbourhoods $\mathcal{A}$ that are all absorbent and balanced.

Again, let U be a neighbourhood of 0 in A. Then as addition C and $0+0 = 0$, there exist open absorbent and balanced (from $\mathcal{A}$) neighbourhoods U_1, U_2 of 0 in A such that $U_1 + U_2 \subseteq U$. □

In Section 4.5 we will see that also demanding convexity of the sets in $\mathcal{A}$ provides topological vector spaces with a lot more structure and will permit us to do analysis (basically, what we are doing is finding a topological basis for A that more and more resembles the topological basis of a metric space which consists of open balls, that are all absorbent, balanced, and convex).

The following examples have been added to emphasise the fact that intuitive results need not be valid in the context of topological vector spaces when we do not place additional constraints on their topologies.

Example 4.3.7: Failure of an almost open mapping to be open
Let A be $\mathbb{R}$ with its usual topology (which makes A FS $/\mathbb{R}$), and let B be $\mathbb{R}$ with topology consisting of $\{\emptyset, \mathbb{R}\}$ (which makes B TVS $/\mathbb{R}$ and Baire).

Then the map $f := \mathrm{id}_{\mathbb{R}} : A \to B$ C 1 $/\mathbb{R}$. Let $U \subseteq A$ be open and nonempty, then there is some $a \in U$, so $B \supseteq \overline{f(U)} \supseteq \overline{\{f(a)\}} = B$ (as $B = \mathbb{R}$ is the smallest closed set containing $f(a)$ in B). Hence $f(U) \subseteq B = \mathrm{int}(\overline{f(U)})$, so f is almost open.

On the other hand, for $U = B_A(0,1)$ we have $f(U) = B_A(0,1) \notin \{\emptyset, \mathbb{R}\}$, so there exists an open $U \subseteq A$ for which $f(U) \subseteq B$ is not open. Hence f is not open.

This shows in particular that the T2 demand in Theorem (4.4.3) is necessary.

Example 4.3.8: Failure of Lemma (3.3.10) for

Consider $f := \mathrm{id}_{\mathbb{R}} : A \to B$ /$\mathbb{R}$ from Example (4.3.7).

Then $\ker f = \{0\}$ and with the usual topology of $\mathbb{R}$, $A = \mathbb{R} \simeq \mathbb{R}/\{0\}$ as (since $U + \{0\} = U$ for all $U \subseteq \mathbb{R}$ open, see the proof of Lemma (4.1.9)). So $A \simeq A/\ker f$, now $B = \mathbb{R} = \mathrm{id}_{\mathbb{R}}(\mathbb{R}) = f(A)$, so by Lemma (3.3.10) we have *as -isomorphisms*

$$A \simeq A/\ker f \simeq f(A) = B.$$

However, these isomorphisms are *not* -isomorphisms, because the topologies on A and B are different.

Hence Lemma (3.3.10) does not hold for general /$\mathbb{K}$.

Note that Lemma (3.3.10) does hold for /$\mathbb{K}$ which satisfy the conditions of Corollary (4.4.6), because the map $A/\ker f \to f(A) : [a] \mapsto f(a)$ /$\mathbb{K}$ and bijective.

Example 4.3.9: Bijective continuous linear map with non-continuous inverse

The map from Example (4.3.7) is /$\mathbb{R}$ and bijective, however, as it is not open, its inverse is not . This shows that not every continuous linear map which is bijective, is a -isomorphism.

In the next section, Corollary (4.4.6) shows that for topological vector spaces with a sufficient amount of structure, the problems in the above examples cannot arise.

4.4 F-spaces

In this section we follow [Hus1965].

Definition 4.4.1: F-space ( )

Let A be a set.

Then we call A an *F-space over* $\mathbb{K}$ (denoted by A /$\mathbb{K}$) if A /$\mathbb{K}$ with a translation invariant metric (that is $d(a_1 + a_3, a_2 + a_3) = d(a_1, a_2)$ for all $a_1, a_2, a_3 \in A$) such that A is complete with respect to this metric.

We will prove the open mapping and closed graph theorems in two stages.

Theorem 4.4.2

Let A, B /$\mathbb{K}$, B Baire (Definition (2.5.16)).

Then

- any $f : A \to B$ /$\mathbb{K}$ that is surjective is almost open,
- any $g : B \to A$ /$\mathbb{K}$ is almost continuous.

Proof.

- Let $f : A \to B$ /$\mathbb{K}$ and surjective. Let U be any open neighbourhood of 0 in A, then by Lemma (4.3.6) there exists an open neighbourhood U_1 of 0 in A that is balanced, $A = \bigcup_{k\in\mathbb{N}} k\,U_1$, and $U_1 + U_1 \subseteq U$.

  Using surjectivity and linearity of f, we therefore find

$$B = f(A) = f\left(\bigcup_{k\in\mathbb{N}} k\,U_1\right) = \bigcup_{k\in\mathbb{N}} f(k\,U_1) = \bigcup_{k\in\mathbb{N}} k\,f(U_1) \subseteq \bigcup_{k\in\mathbb{N}} k\,\overline{f(U_1)} \subseteq B.$$

Because B is Baire and has nonempty interior itself (being B), there exists a $k \in \mathbb{N}$ such that $\mathrm{int}(k\,\overline{f(U_1)}) \neq \emptyset$, therefore (Lemma (4.1.5)) $\mathrm{int}(\overline{f(U_1)}) \neq \emptyset$. Hence there exists a $b \in \mathrm{int}(\overline{f(U_1)})$.

Since U_1 is balanced and f [1], by Lemma (4.3.5) $f(U_1)$, $\overline{f(U_1)}$ are balanced. Therefore (again Lemma (4.3.5)) $\mathrm{int}(\overline{f(U_1)}) = -\mathrm{int}(\overline{f(U_1)})$. Hence $0 = b-b \in \mathrm{int}(\overline{f(U_1)}) - \mathrm{int}(\overline{f(U_1)}) = \mathrm{int}(\overline{f(U_1)}) + \mathrm{int}(\overline{f(U_1)}) \subseteq \mathrm{int}(\overline{f(U_1)} + \overline{f(U_1)}) \subseteq \mathrm{int}(\overline{f(U_1) + f(U_1)}) = \mathrm{int}(\overline{f(U_1 + U_1)}) \subseteq \mathrm{int}(\overline{f(U)})$ (using continuity of addition for Lemma (2.1.16), and linearity of f). Hence $f(0) = 0 \in \mathrm{int}(\overline{f(U)})$ for any open neighbourhood of 0 in A.

As f [1] we therefore find that f is almost open.

- Let $g : B \to A$ [1] $/\mathbb{K}$. Let U be any open neighbourhood of $g(0) = 0$ in A, then by Lemma (4.3.6) there exists an open neighbourhood U_1 of 0 in A that is balanced, $A = \bigcup_{k\in\mathbb{N}} k\,U_1$, and $U_1 + U_1 \subseteq U$. Then

  $$B = g^{-1}(A) = g^{-1}\left(\bigcup_{k\in\mathbb{N}} k\,U_1\right) = \bigcup_{k\in\mathbb{N}} k\,g^{-1}(U_1) \subseteq \bigcup_{k\in\mathbb{N}} k\,\overline{g^{-1}(U_1)} \subseteq B.$$

  So as B is Baire, there exists a $b \in \mathrm{int}(\overline{g^{-1}(U_1)})$. Using Lemma (4.3.5) we see that $\mathrm{int}(U_1) = -\mathrm{int}(U_1)$, so $0 = b - b \in \mathrm{int}(\overline{g^{-1}(U_1)}) - \mathrm{int}(\overline{g^{-1}(U_1)}) = \mathrm{int}(\overline{g^{-1}(U_1)}) + \mathrm{int}(\overline{g^{-1}(U_1)}) \subseteq \mathrm{int}(\overline{g^{-1}(U_1)} + \overline{g^{-1}(U_1)}) \subseteq \mathrm{int}(\overline{g^{-1}(U_1 + U_1)}) \subseteq \mathrm{int}(\overline{g^{-1}(U)})$.

  From linearity of g and the fact that $0 \in \mathrm{int}(\overline{g^{-1}(U)}) \subseteq \overline{g^{-1}(U)}$ for all open neighbourhoods U of $g(0)$ in A we therefore find that g is almost continuous.

□

Theorem 4.4.3: Open mapping theorem
Let A [FS] $/\mathbb{K}$, B [Vs] $/\mathbb{K}$ [T2].

Then any $f : A \to B$ [1] $/\mathbb{K}$ that is almost open and [C], is in fact open.

Proof. Let $U \subseteq A$ be open, then by Lemma (4.1.6) and Definition (4.5.1) we have that for any $a \in U$ there exists a neighbourhood U_a of 0 in A such that $a + U_a \subseteq U$, using linearity of f we therefore find that

$$f(U) \supseteq f\left(\bigcup_{a\in U}(a + U_a)\right) = \bigcup_{a\in U}(f(a) + f(U_a)) \supseteq \bigcup_{a\in U}(f(a) + \{0\}) = f(U).$$

Hence, if for all neighbourhoods U_a of 0 in A we can find an open neighbourhood V_a of 0 in B such that $\{0\} \subseteq V_a \subseteq f(U_a)$, we would obtain that

$$f(U) = \bigcup_{a\in U}(f(a) + V_a)$$

which is open by Lemma (4.1.5). Since U was arbitrary, this would imply that f is open.

To prove the theorem it is sufficient to show that for any neighbourhood U of 0 in A there exists an open neighbourhood V of 0 in B such that $V \subseteq f(U)$. Conversely, any open map has this property.

Construct a collection $U_k := \overline{B_A(0, 2^{-k})} \subseteq A$ for all $k \in \mathbb{N}$, then by definition $U_1 \supseteq U_2 \supseteq \ldots$. Let $a_1, a_2 \in U_{k+1}$, then $d(a_1 + a_2, 0) = d(a_1, -a_2) \leq d(a_1, 0) + d(0, -a_2) = d(a_1, 0) + d(a_2, 0) \leq 2^{-(k+1)} + 2^{-(k+1)} = 2^{-k}$, so $a_1 + a_2 \in U_k$. Therefore $U_k \supseteq U_{k+1} + U_{k+1}$ for all $k \in \mathbb{N}$.

Note that for any $k \in \mathbb{N}$, $0 \in B_A(0, 2^{-k}) \subseteq U_k$, so U_k is a closed neighbourhood of 0 in A. Let U be any open neighbourhood of 0 in A, then (A d(.,.)), there exists an $\epsilon \in]0, \infty[$ such that $B_A(0, \epsilon) \subseteq U$. Therefore, for $k = \lceil {}^2\log(1/\epsilon) \rceil + 1 \in \mathbb{N}$ we have $2^{-k} < \epsilon$, so $0 \in U_k \subseteq B_A(0, \epsilon) \subseteq U$. Hence $U_1, U_2, \ldots$ is a basis of closed neighbourhoods of 0 in A.

Define for all $k \in \mathbb{N}$, $V_k := \operatorname{int}(\overline{f(U_k)})$. Because f is almost open and linear, $0 = f(0) \in f(U_k) \subseteq \operatorname{int}(\overline{f(U_k)}) = V_k \subseteq \overline{f(U_k)}$, so all V_k are open neighbourhoods of 0 in B and $0 \in V_k \subseteq \overline{f(U_k)}$ for all $k \in \mathbb{N}$.

To prove the theorem it is therefore sufficient to show that $V_{k+1} \subseteq f(U_k)$ for all $k \in \mathbb{N}$, because $U_1, U_2, \ldots$ forms a basis of neighbourhoods of 0 in A.

Fix $k \in \mathbb{N}$ and $b \in V_{k+1}$. We are going to inductively create a sequence in U_k that approximates b.

As $b \in V_{k+1} \subseteq \overline{f(U_{k+1})}$ and $b - V_{k+2}$ is an open neighbourhood of b in B, by Lemma (2.1.3) $f(U_{k+1}) \cap (b - V_{k+2}) \neq \emptyset$: there is some $a_1 \in U_{k+1}$ with $b - f(a_1) \in V_{k+2}$.

Now suppose that we have constructed $a_1, \ldots, a_l$ such that $b - f\left(\sum_{m=1}^{l} a_m\right) \in V_{k+1+l}$ and $a_m \in U_{k+m}$ for all $1 \leq m \leq l$. Let $b' := b - f(a_1 + \ldots + a_l) \in V_{k+1+l} \subseteq \overline{f(U_{k+1+l})}$. Then $b' - V_{k+1+l+1}$ is an open neighbourhood of b' in B, so $f(U_{k+1+l}) \cap (b' - V_{k+1+l+1}) \neq \emptyset$. This implies that there exists a $a_{l+1} \in U_{k+1+l}$ with $b - f(a_1 + \ldots + a_l + a_{l+1}) = (b - f(a_1 + \ldots + a_l)) - f(a_{l+1}) = b' - f(a_{l+1}) \in V_{k+1+l+1}$.

Using induction this permits us to construct a sequence $\mathbb{N} \to A : l \mapsto a_l$ satisfying for all $l \in \mathbb{N}$ that

$$a_l \in U_{k+l}, \qquad b - f\left(\sum_{m=1}^{l} a_m\right) \in V_{k+l+1}. \tag{4.1}$$

Let $l, m \in \mathbb{N}$, then because $U_1 \supseteq U_2 + U_2 \supseteq \ldots$ we have

$$\begin{aligned}
\sum_{n=1}^{l} a_{n+m} &= a_{m+1} + a_{m+2} + \ldots + a_{m+l-2} + a_{m+l-1} + a_{m+l} \\
&\in U_{k+m+1} + U_{k+m+2} + \ldots + U_{k+m+l-2} + U_{k+m+l-1} + U_{k+m+l} \\
&\subseteq U_{k+m+1} + U_{k+m+2} + \ldots + U_{k+m+l-2} + U_{k+m+l-1} + U_{k+m+l-1} \\
&\subseteq U_{k+m+1} + U_{k+m+2} + \ldots + U_{k+m+l-2} + U_{k+m+l-2} \\
&\ldots \\
&\subseteq U_{k+m+1} + U_{k+m+1} \\
&\subseteq U_{k+m}.
\end{aligned}$$

So

$$\sum_{n=1}^{l} a_{n+m} \in U_{k+m}$$

for any $l, m \in \mathbb{N}$.

Let $\epsilon \in]0, \infty[$ be given and pick $l = \lceil {}^2\log(1/\epsilon) \rceil + 1 \in \mathbb{N}$, then $2^l < \epsilon$. Let $m, n \geq l$, $m \leq n$. Then

$$\sum_{o=1}^{n} a_o - \sum_{p=1}^{m} a_p = \sum_{o=1}^{n-m} a_{o+m} \in U_{k+m} \subseteq U_{k+l}$$

and hence $d(\sum_{o=1}^{n} a_o, \sum_{p=1}^{m} a_p) \leq 2^{-(k+l)} \leq 2^{-l} < \epsilon$.

Therefore the sequence $l \mapsto \sum_{m=1}^{l} a_m$ is Cauchy and because A is complete, there exists an $a \in A$ such that (recall that U_k is closed, use Lemma (2.1.16))

$$a = \lim_{l\to\infty} \sum_{m=1}^{l} a_m = \sum_{l=1}^{\infty} a_{l+0} \in \overline{U_{k+0}} = U_k.$$

As f C we find for this a that

$$b - f(a) = b - f\left(\lim_{l\to\infty} \sum_{m=1}^{l} a_m\right) = \lim_{l\to\infty}\left(b - f\left(\sum_{m=1}^{l} a_m\right)\right).$$

By Equation (4.1) we see that for any $l \in \mathbb{N}$ we have that for all $m \in \mathbb{N}$

$$b - f\left(\sum_{n=1}^{l+m} a_n\right) \in V_{k+l+1},$$

so taking the limit $m \to \infty$ we obtain that for any $l \in \mathbb{N}$, $b - f(a) \in \overline{V_{k+l+1}}$. Hence

$$b - f(a) \in \bigcap_{l\in\mathbb{N}} \overline{f(U_{k+l+1})}.$$

Let $b_1 \in \bigcap_{l\in\mathbb{N}} \overline{f(U_{k+l+1})}$ be arbitrary. Let V be an arbitrary open neighbourhood of 0 in B. Then for all $l \in \mathbb{N}$, $f(U_{k+l+1}) \cap (b_1 + V) \neq \emptyset$, so we obtain for each $l \in \mathbb{N}$ an $a'_l \in U_{k+l+1}$ such that $f(a'_l) \in b_1 + V$. Because the U_{k+l+1} form a decreasing basis of neighbourhoods of 0 in A, $\lim_{l\to\infty} a'_l = 0$. Hence ($f$ C) $\lim_{l\to\infty} f(a'_l) = 0$, so $-b_1 = \lim_{l\to\infty}(f(a'_l) - b_1) \in \overline{V}$. So $-b_1 \in \overline{V}$ for any open neighbourhood V of 0 in B. Suppose $-b_1 \neq 0$, then (B T2) there exists an open neighbourhood V of 0 in B and V_1 of $-b_1$ in B such that $V \cap V_1 = \emptyset$. By the above $-b_1 \in \overline{V} \subseteq A \setminus V_1$ (as $A \setminus V_1$ closed and $V \subseteq A \setminus V_1$), so $-b_1 \notin V_1$. This leads to a contradiction with the fact that V_1 is an open neighbourhood of $-b_1$. Hence $-b_1 = 0$ and therefore $b_1 = 0$.

On the other hand, for all $l \in \mathbb{N}$, $0 = f(0) \in f(U_{k+l+1}) \subseteq \overline{f(U_{k+l+1})}$, so

$$\bigcap_{l\in\mathbb{N}} \overline{f(U_{k+l+1})} = \{0\}.$$

Therefore $b - f(a) = 0$ and $b = f(a)$.

Hence, for any $b \in V_{k+1}$ there exists an $a \in U_k$ such that $f(a) = b$, hence $V_{k+1} \subseteq f(U_k)$ for all $k \in \mathbb{N}$ and the theorem is proven. □

Theorem 4.4.4: Closed graph theorem

Let A FS $/\mathbb{K}$, B TVS $/\mathbb{K}$.

Then any $g : B \to A$ l $/\mathbb{K}$ that is almost continuous and for which $\operatorname{graph}(g) \subseteq B \times A$ is closed, is in fact continuous.

Proof. As g $\textcircled{\scriptsize I}$ it is sufficient to show that g $\textcircled{\scriptsize c}$ 0.

We are now going to show that we may assume g to be injective without loss of generality.

Let $f : B \to B/\ker g : b \mapsto [b]$. Then g factorises uniquely (Lemma (3.3.10)) as a map $g = h \circ f$ with $h : B/\ker g \to A : [b] \mapsto g(b)$. g is almost continuous by assumption, so for any $b \in B$ and open neighbourhood U of $g(b)$ in A we have that $\overline{g^{-1}(U)}$ is a neighbourhood of b. As $g^{-1}(U) = f^{-1}(h^{-1}(U))$ and f $\textcircled{\scriptsize c}$, $\overline{h^{-1}(U)} = \overline{f(g^{-1}(U))} \supseteq f(\overline{g^{-1}(U)})$. Because f is open (Lemma (4.1.9)), this implies that $\overline{h^{-1}(U)}$ is a neighbourhood of $[b]$ in $B/\ker g$. This shows that h is almost continuous.

Now if h $\textcircled{\scriptsize c}$, then $g = h \circ f$ $\textcircled{\scriptsize c}$. Therefore it is sufficient to prove the theorem for h, which is injective by Lemma (3.3.10), and we may assume g to be injective.

Construct the same countable basis of closed neighbourhoods of 0 in A as in Theorem (4.4.3): $U_1 \supseteq U_2 \supseteq \ldots$ satisfying $U_k \supseteq U_{k+1} + U_{k+1}$ for each $k \in \mathbb{N}$.

Define for all $k \in \mathbb{N}$, $V_k := \operatorname{int}(\overline{g^{-1}(U_k)})$. Because g is almost continuous, $\overline{g^{-1}(U_k)}$ is a neighbourhood of $0 = g(0)$ in B, so $0 \in V_k \subseteq \overline{g^{-1}(U_k)}$ for each $k \in \mathbb{N}$.

To prove the theorem it is therefore sufficient to show that $g(V_{k+1}) \subseteq U_k$ for each $k \in \mathbb{N}$ (as the U_k form a basis of closed neighbourhoods of 0 in A and all V_k are open neighbourhoods of 0 in B).

Fix $k \in \mathbb{N}$ and $b \in V_{k+1}$. We are going to create a sequence in A that approximates $g(b)$.

As $b \in V_{k+1} \subseteq \overline{g^{-1}(U_{k+1})}$ and $b - V_{k+2}$ is an open neighbourhood of b in B by Lemma (4.1.5), $g^{-1}(U_{k+1}) \cap (b - V_{k+2}) \neq \emptyset$. So there exists some $b_1 \in g^{-1}(U_{k+1}) \cap (b - V_{k+2})$, hence $b - b_1 \in V_{k+2}$ and $g(b_1) \in U_{k+1}$.

Now suppose we have constructed $b_1, \ldots, b_l$ such that $g(b_m) \in U_{k+m}$ for $1 \le m \le l$ and $b - \sum_{m=1}^{l} b_m \in V_{k+l+1}$. Let $b' := b - (b_1 + \ldots + b_l) \in V_{k+l+1} \subseteq \overline{g^{-1}(U_{k+l+1})}$, then $b' - V_{k+1+l+1}$ is an open neighbourhood of b' in B, so there exists some $b_{l+1} \in g^{-1}(U_{k+l+1}) \cap (b' - V_{k+1+l+1})$. Hence $b - (b_1 + \ldots + b_l + b_{l+1}) = b' - b_{l+1} \in V_{k+1+l+1}$ and $g(b_{l+1}) \in U_{k+l+1}$.

Using induction this permits us to construct a sequence $\mathbb{N} \to B : l \mapsto b_l$, satisfying for all $l \in \mathbb{N}$ that

$$g(b_l) \in U_{k+l}, \qquad b - \sum_{m=1}^{l} b_m \in V_{k+l+1}.$$

As $U_1 \supseteq U_2 + U_2 \supseteq \ldots$ we find via the same reasoning as in Theorem (4.4.3) that $l \mapsto \sum_{m=1}^{l} g(b_l)$ is a Cauchy sequence in A and because A is complete there exists an $a \in A$ such that

$$a = \sum_{l=1}^{\infty} g(b_l) \in \overline{U_k} = U_k.$$

Now as for any $l, m \in \mathbb{N}$

$$b - \sum_{n=1}^{l+m} b_n \in V_{k+l+m+1} \subseteq V_{k+l+1}$$

we find taking the limit of $m \to \infty$ that for all $l \in \mathbb{N}$, $b - \sum_{m=1}^{\infty} b_m \in \overline{V_{k+l+1}} = \overline{g^{-1}(U_{k+l+1})}$. Hence

$$b - \sum_{m=1}^{\infty} b_m \in \bigcap_{l \in \mathbb{N}} \overline{g^{-1}(U_{k+l+1})}.$$

Let $b_1 \in \bigcap_{l\in\mathbb{N}} \overline{g^{-1}(U_{k+l+1})}$. Let $V \times U$ any open neighbourhood of $(0,0)$ in $B \times A$. Then for all $l \in \mathbb{N}$, $b_1 \in \overline{g^{-1}(U_{k+l+1})} \cap (b_1 + V)$, so there exists an $b'_l \in b_1 + V$ with $g(b'_l) \in U_{k+l+1}$. As the $\{U_{k+l+1} | l \in \mathbb{N}\}$ form a descending basis of neighbourhoods of 0 in A, there exists an $l \in \mathbb{N}$ such that $U_{k+l+1} \subseteq U$. Hence $(b'_l, g(b'_l)) \in (b_1+V)\times U \cap \operatorname{graph}(g)$. So $(b_1+V)\times(0+U)\cap\operatorname{graph}(g) \neq \emptyset$ for all open neighbourhoods $V \times U$ of $(0,0)$ in $B \times A$. Hence as $\operatorname{graph}(g)$ is closed, $(b_1, 0) \in \operatorname{graph}(g)$, so $g(b_1) = 0$. On the other hand, for all $l \in \mathbb{N}$, $0 \in g^{-1}(\{0\}) \subseteq \overline{g^{-1}(U_{k+l+1})}$, so

$$\{0\} \subseteq \bigcap_{l\in\mathbb{N}} \overline{g^{-1}(U_{k+l+1})} \subseteq \ker g.$$

As g is injective, $\ker g = \{0\}$ and therefore $b - \sum_{m=1}^{\infty} b_m = 0$.

Now consider the sequence $\mathbb{N} \to \operatorname{graph}(g) : l \mapsto (b - \sum_{m=1}^{l} b_m, g(b - \sum_{m=1}^{l} b_m)) = (b - \sum_{m=1}^{l} b_m, g(b) - \sum_{m=1}^{l} g(b_m))$ which as $l \to \infty$ goes to $(0, g(b) - a) \in \operatorname{graph}(g)$ as $\operatorname{graph}(g)$ is closed. Hence $g(b) - a = g(0) = 0$, so $g(b) = a \in U_k$ and therefore $g(V_{k+1}) \subseteq U_k$ for all $k \in \mathbb{N}$ and the theorem is proven. □

We can now combine the above results with Theorem (4.4.2).

Theorem 4.4.5: Banach
Let A FS $/\mathbb{K}$, B TVS $/\mathbb{K}$ T2 Baire.
Then

- any $f : A \to B$ C 1 $/\mathbb{K}$ is surjective if and only if f is open,
- any $g : B \to A$ 1 $/\mathbb{K}$ is continuous if and only if $\operatorname{graph}(g) \subseteq B \times A$ is closed.

Proof.
- Let $f : A \to B$ C 1 $/\mathbb{K}$. Suppose f is surjective, then by Theorem (4.4.2) (B Baire), f is almost open. Hence by Theorem (4.4.3) (B T2), f is open.

  Conversely, suppose f is open. Then for any open neighbourhood U of 0 in A, $f(U)$ is an open neighbourhood of 0 in B. By Lemma (4.3.6), U and $f(U)$ are absorbent. Since f 1 we therefore obtain that $B = \bigcup_{\alpha\in]0,\infty[} \alpha f(U) = \bigcup_{\alpha\in]0,\infty[} f(\alpha U) = f\left(\bigcup_{\alpha\in]0,\infty[} \alpha U\right) = f(A)$. Therefore f is surjective.

- Let $g : B \to A$ 1 $/\mathbb{K}$. Suppose $\operatorname{graph}(g)$ is closed, then by Theorem (4.4.2), g is almost continuous. Hence by Theorem (4.4.4), g C.

  Conversely, if g C, by Lemma (2.2.7), $\operatorname{graph}(g)$ is closed since A T2 by Theorem (2.5.10).

□

Corollary 4.4.6
Let A FS $/\mathbb{K}$, B TVS $/\mathbb{K}$ T2 Baire, $f : A \to B$.

If f C I $/\mathbb{K}$ and bijective, then its inverse is C I $/\mathbb{K}$ and bijective.

If this is the case, then B is also FS $/\mathbb{K}$.

Proof. Suppose f C I $/\mathbb{K}$ and bijective. Because f is bijective and I, there exists an inverse $g : B \to A$ which is bijective and I. B is Baire and f is surjective because f is bijective, so by Theorem (4.4.2) f is almost open. A FS and B T2, so by Theorem (4.4.3) f is open. In particular, for any $U \subseteq A$ open, $g^{-1}(U) = f(U) \subseteq B$ is open (g is bijective with inverse f), so g C by Lemma (2.1.14).

This makes B T-isomorphic to A and hence metrisable with a translation invariant (f and g are linear) metric and complete (use f and g to move Cauchy sequences between A and B). Hence B FS $/\mathbb{K}$. □

4.5 Local convexity

Definition 4.5.1: Locally convex topological vector spaces (LC)
Let A TVS $/\mathbb{K}$.

Then we call A *locally convex* (denoted by A LC) if there exists a basis of open neighbourhoods $\mathcal{A}$ of 0 in A such that all $U \in \mathcal{A}$ are abc subsets of A.

Lemma 4.5.2
Let A TVS $/\mathbb{K}$.

Then for any open neighbourhood U of 0 in A that is an abc subset of A, the map
$$\|\cdot\|_U : A \to \mathbb{R} : a \mapsto \inf\{\alpha \in]0,\infty[\mid a \in \alpha\, U\}$$
is a seminorm on A and for all $\alpha \in]0,\infty[$,
$$\alpha\, U = \|\cdot\|_U^{-1}([0,\alpha[).$$

Conversely, for any seminorm $\|\cdot\| : A \to \mathbb{R}$, the set $\|\cdot\|^{-1}([0,1[)$ is an abc subset of A.

Proof. By definition $\|a\|_U \geq 0$ for all $a \in A$ and since U is absorbent we have that $\forall a \in A : \exists \alpha \in]0,\infty[: a \in \alpha\, U$, so $\|a\|_U \in [0,\alpha] \subseteq \mathbb{R}$ exists for all $a \in A$. For $\alpha = 0$ we clearly have $\|\alpha\, a\|_U = \inf]0,\infty[= 0 = 0\|a\|_U$, if $\alpha \neq 0$ we see that for any $\beta \in]0,\infty[$, $\alpha\, a \in \beta\, U$ iff $\alpha\, a = \beta\, a_1$ for some $a_1 \in U$ iff $a = \frac{\beta}{|\alpha|}\, \frac{|\alpha|}{\alpha}\, a_1$ for some $a_1 \in U$ iff $a = \frac{\beta}{|\alpha|}\, a_2$ for some $a_2 \in U$ (because U is balanced and $||\alpha|/\alpha| = 1$) iff $a \in \frac{\beta}{|\alpha|}\, U$. Therefore $\|\alpha\, a\|_U = \inf\{\beta \in]0,\infty[| \alpha\, a \in \beta\, U\} = \inf\{\beta \in]0,\infty[| a \in \frac{\beta}{|\alpha|}\, U\} = |\alpha|\, \|a\|_U$. So $\|\alpha a\|_U = |\alpha|\, \|a\|_U$ for all $a \in A$, $\alpha \in \mathbb{K}$. Let $a_1 \in \alpha_1 U$ and $a_2 \in \alpha_2 U$ for $\alpha_1, \alpha_2 \in]0,\infty[$, then there exist $a_3, a_4 \in U$ such that $a_1 = \alpha_1\, a_3$, $a_2 = \alpha_2\, a_4$. Because of this $a_1 + a_2 = \alpha_1\, a_3 + \alpha_2\, a_4 = (\alpha_1 + \alpha_2)((\alpha_1/(\alpha_1+\alpha_2))\, a_3 + (\alpha_2/(\alpha_1+\alpha_2))\, a_4) \in (\alpha_1+\alpha_2)\, U$, as U is convex and $(\alpha_1/(\alpha_1+\alpha_2)) + (\alpha_2/(\alpha_1+\alpha_2)) = 1$. Therefore $\|a_1+a_2\|_U \leq \|a_1\|_U + \|a_2\|_U$. This makes $\|\cdot\|_U$ a seminorm on A.

Now let $\alpha \in]0,\infty[$ be given. Let $a \in A$ and suppose $\|a\|_U < \alpha$, then there exists a $\beta \in [0,\alpha[$ such that $a \in \beta\, U \subseteq \alpha\, U$ by definition of the infimum. Therefore $\|\cdot\|_U^{-1}([0,\alpha[) \subseteq \alpha\, U$. Let $a \in \alpha U$, then $a = \alpha\, a_1$ for some $a_1 \in U$. Because the map $\mathbb{R} \to A : \beta \mapsto \beta a_1$ C and $\alpha\, U$ is an open neighbourhood of $\alpha\, a_1$ in A there exists a $\delta \in]0, \infty[$ such that $\beta\, a_1 \in \alpha U$ for all $\beta \in]\alpha(1-\delta), \alpha(1+\delta)[$. But then $(1+\delta/2)\, a = \alpha\,(1+\delta/2)\, a_1 \in U$, so $a \in (\alpha/(1+\delta/2))\, U$ and $\|a\|_U \leq \alpha/(1+\delta/2) < \alpha$. Therefore $\alpha\, U \subseteq \|\cdot\|_U^{-1}([0, \alpha[)$.

Let $\|\cdot\| : A \to \mathbb{R}$ be a seminorm and define $U := \|\cdot\|^{-1}([0,1[) \subseteq A$.

Let $a \in A$. Suppose $\|a\| = 0$, then $a \in U$ directly. Otherwise $\|a\| > 0$ (as $\|a\| \geq 0$), so we can define $a_1 := \frac{1}{2\,\|a\|}\, a$. Then $\|a_1\| = \frac{1}{2} \in [0, 1[$, so $a_1 \in U$ and therefore $a \in 2\,\|a\|\, U$. Hence U is absorbent.

Let $a \in U$, $\alpha \in \overline{B_{\mathbb{K}}(0,1)}$, then $0 \leq \|\alpha\, a\| = |\alpha|\, \|a\| \leq 1\, \|a\| < 1$, so $\alpha\, a \in U$: U is balanced.

Let $a_1, a_2 \in U$, $\alpha \in [0,1]$. Then $0 \leq \|\alpha\, a_1 + (1-\alpha)\, a_2\| \leq |\alpha|\, \|a_1\| + |1-\alpha|\, \|a_2\| < \alpha\, 1 + (1-\alpha)\, 1 = 1$, so $\alpha\, a_1 + (1-\alpha)\, a_2 \in U$. Hence U is convex.

So U is an abc subset of A. □

Lemma 4.5.3

Let A Vs $/\mathbb{K}$.

Then A LC if and only if there exists a family of seminorms $\{\|\cdot\|_i : A \to \mathbb{R} | i \in I\}$ on A such that the topology of A is the initial topology (Definition (2.1.18)) with respect to this collection.

Proof. Suppose A LC and let $\mathcal{A}$ be the basis of abc neighbourhoods of 0 in A. Define for each $U \in \mathcal{A}$ the seminorm $\|\cdot\|_U : A \to \mathbb{R}$ as in Lemma (4.5.2). With Lemma (4.5.2) and Lemma (4.1.6) we obtain that the initial topology of $\{\|\cdot\|_U | U \in \mathcal{A}\}$ must coincide with the topology of A generated by the basis of neighbourhoods $\mathcal{A}$: for any $U_1, \ldots, U_k \in \mathcal{A}$ and $\epsilon \in]0, \infty[$ we have $\|\cdot\|_{U_1}^{-1}([0,\epsilon[) \cap \ldots \cap \|\cdot\|_{U_k}^{-1}([0,\epsilon[) = (\epsilon\, U_1) \cap \ldots \cap (\epsilon\, U_k) = \epsilon\,(U_1 \cap \ldots \cap U_k)$.

Suppose conversely that the topology of A is the initial topology of a family of seminorms $\{\|\cdot\|_i : A \to \mathbb{R} | i \in I\}$. Choose

$$\begin{aligned} \mathcal{A} := \{\, \{a \in A \mid \|a\|_{i_1}, \ldots, \|a\|_{i_k} < \epsilon\} \mid i_1, \ldots, i_k \in I, \epsilon \in]0, \infty[\} \\ = \{\|\cdot\|_{i_1}^{-1}(]-1, \epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1, \epsilon[) \\ \mid i_1, \ldots, i_k \in I, \epsilon \in]0, \infty[\}. \end{aligned}$$

Then all these sets are open, because all sets $]-1, \epsilon[\subseteq \mathbb{R}$ are open for $\epsilon \in]0, \infty[$ and the $\|\cdot\|_i$ C by choice of the initial topology. Furthermore 0 is an element of all these sets, since $\|0\|_i = 0 < \epsilon$ for all $\epsilon \in]0, \infty[$ and $i \in I$ because the $\|\cdot\|_i$ are seminorms. Therefore $\mathcal{A}$ is a collection of open neighbourhoods of 0 in A. From the expression for the basis generating the initial topology (Lemma (2.1.19)) we see that it is even a basis of open neighbourhoods of 0 (which by Lemma (4.1.6) generates the entire initial topology by translation). The sets are furthermore all abc subsets of A by Lemma (4.5.2). Therefore A LC. □

For the remaining part of this section we will use the notion of local convexity LC interchangeably with the family of seminorms $\{\|\cdot\|_i | i \in I\}$ from Lemma (4.5.3).

Lemma 4.5.4

Let A Vs $/\mathbb{K}$ LC, B T.

Then for $f : B \to A$, $b \in B$, $a \in A$, the following are equivalent:

- $\lim_{y\to b} f(y) = a$,
- for all $i \in I$, $\lim_{y\to b} \|f(y) - a\|_i = 0$,
- $\forall i \in I : \forall \epsilon \in]0, \infty[: \exists b \in V \subseteq B$ open $: \forall b_1 \in V : \|f(b_1) - a\|_i < \epsilon$.

Proof. By considering $b \mapsto f(b) - a$ with continuity of addition we may assume $a = 0$.

Suppose $\lim_{y\to b} f(y) = 0$. Since A has the initial topology of all $\|\cdot\|_i$, all $\|\cdot\|_i$ and hence by Lemma (2.1.10), $\lim_{y\to b} \|f(y)\|_i = \lim_{x\to 0} \|x\|_i = \|0\|_i = 0$ for all $i \in I$

Suppose conversely that $\lim_{y\to b} \|f(y)\|_i = 0$ for all $i \in I$. Let U be an open neighbourhood of 0 in A, then because of the initial topology there exists an $\epsilon \in]0, \infty[$ and $i_1, \ldots, i_k \in I$ such that $0 \in \|\cdot\|_{i_1}^{-1}(]-1, \epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1, \epsilon[) \subseteq U$. For each $i_j \in I$ with $1 \leq j \leq k$ there exists an open neighbourhood V_j of b in B such that $\|f(V_j)\|_{i_j} \subseteq]-1, \epsilon[$ since $\lim_{y\to b} \|f(y)\|_{i_j} = 0$ by assumption and $]-1, \epsilon[$ is an open neighbourhood of 0 in $\mathbb{R}$. Therefore, $f(V_j) \subseteq \|\cdot\|_{i_j}^{-1}(]-1, \epsilon[)$ for each $1 \leq j \leq k$. As a finite intersection of open neighbourhoods, $V := V_1 \cap \ldots \cap V_k$ is an open neighbourhood of b in B and by construction $f(V) \subseteq \|\cdot\|_{i_1}^{-1}(]-1, \epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1, \epsilon[) \subseteq U$. So for each open neighbourhood U of 0 in A there exists an open neighbourhood V of b in B such that $f(V) \subseteq U$. Therefore $\lim_{y\to b} f(y) = 0$.

The final statement arises from writing out the second explicitly. □

Lemma 4.5.5
Let A, B $/\mathbb{K}$, $f : A \to B$ $/\mathbb{K}$.

Denote the seminorms on A by $\|\cdot\|_i$ for $i \in I$ and the seminorms on B by $\|\cdot\|'_j$ for $j \in J$.

Then f if and only if for all $j \in J$ there exist $\alpha \in]0, \infty[$ and $i_1, \ldots, i_k \in I$ with
$$\|f(a)\|'_j \leq \alpha\Big(\|a\|_{i_1} + \ldots + \|a\|_{i_k}\Big)$$
for all $a \in A$.

Proof. We follow [Bou1955]. Because f and Lemma (4.1.5) we know that f if and only if f 0.

Suppose the estimate holds and let $\epsilon \in]0, \infty[$, $j \in J$. Then there exist $\alpha \in]0, \infty[$ and $i_1, \ldots, i_k \in I$ such that the estimate holds. Pick $\delta := \frac{\epsilon}{k\alpha} \in]0, \infty[$, then for all $a \in A$ with $\|a\|_{i_1} \leq \delta, \ldots, \|a\|_{i_k} \leq \delta$ we have $\|f(a)\|'_j \leq \alpha k \frac{\epsilon}{k\alpha} = \epsilon$. Hence $\lim_{a\to 0} \|f(a)\|'_j = 0$ for all $j \in J$ and therefore (Lemma (4.5.4)) $\lim_{a\to 0} f(a) = 0$ which makes f .

Suppose conversely that f , then $\lim_{a\to 0} f(a) = 0$, so $f^{-1}(\|\cdot\|_j'^{-1}(]-1, 1[))$ is a neighbourhood of 0 in A. Hence (by Lemma (2.1.19)) there exist $i_1, \ldots, i_k \in I$ and an $\epsilon \in]0, \infty[$ such that $0 \in \|\cdot\|_{i_1}^{-1}(]-1, \epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1, \epsilon[) \subseteq f^{-1}(\|\cdot\|_j'^{-1}(]-1, 1[))$. Let $a \in A$ be arbitrary and suppose we have some $\gamma \in]0, \infty[$ such that $\gamma\|a\|_{i_l} < \epsilon$ for all $1 \leq l \leq k$. Then by construction of ϵ, $\|f(\gamma a)\|'_j < 1$, so $0 \leq \|f(a)\|'_j < \frac{1}{\gamma}$. Now if $\|a\|_{i_l} = 0$ for some $1 \leq l \leq k$, we can let $\gamma \to \infty$ which shows that $\|f(a)\|'_j = 0$ and the estimate holds. Otherwise pick $\gamma = \frac{\epsilon}{\|a\|_{i_1} + \ldots + \|a\|_{i_k}}$ from which we find
$$\|f(a)\|'_j \leq \frac{1}{\gamma} = \frac{1}{\epsilon}(\|a\|_{i_1} + \ldots + \|a\|_{i_k}),$$

which shows that the desired estimate holds for all $a \in A$ if we pick $\alpha = \frac{1}{\epsilon}$. ☐

Lemma 4.5.6
Let A Vs $/\mathbb{K}$ LC.

The A T2 if and only if for all $a \in A$ we have that $a = 0$ if $\|a\|_i = 0$ for all $i \in I$.

Proof. Suppose A T2. Let $a \in A$ and suppose $a \neq 0$. Since A T2 and the initial topology is generated by the seminorms there exist an $\epsilon \in]0, \infty[$ and $i_1, \ldots, i_k \in I$ such that $a \notin \|\cdot\|_{i_1}^{-1}(]-1, \epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1, \epsilon[) \ni 0$. Therefore, for some $l \in \{1, \ldots, k\}$ we have $a \notin \|\cdot\|_{i_l}^{-1}(]-1, \epsilon[)$, so $\|a\|_{i_l} \geq \epsilon > 0$. So for all $a \in A$, if $a \neq 0$, there exists an $i \in I$ such that $\|a\|_i \neq 0$. Now take the contrapositive.

Suppose for all $a \in A$, if for all $i \in I$ we have $\|a\|_i = 0$, then $a = 0$. Let $a_1, a_2 \in A$ be given and suppose $a_1 \neq a_2$. By translating by $-a_1$ we may suppose that $a_1 = 0$, $a_2 = a \neq 0$. Since $a \neq 0$, there exists an $i \in I$ such that $\|a\|_i > 0$ by assumption. Take for $\epsilon := \|a\|_i/2 > 0$ the sets $U_1 := \|\cdot\|_i^{-1}(]-1, \epsilon[)$ of 0 and $U_2 := a + \|\cdot\|_i^{-1}(]-1, \epsilon[)$. It is clear that U_1 resp. U_2 is an open neighbourhood of 0 resp. a in A. Let $a_1 \in U_1 \cap U_2$, then because $a_1 \in U_1$ we have $\|a_1\|_i < \epsilon$ and because $a_1 \in U_2$, $a_1 = a + a_2$ with $\|a_2\|_i < \epsilon$. But then $2\,\epsilon = \|a\|_i = \|a_1 - a_2\|_i \leq \|a_1\|_i + \|a_2\|_i < 2\epsilon$ leading to a contradiction. Therefore $U_1 \cap U_2 = \emptyset$, so U_1 and U_2 are two disjoint open neighbourhoods. This makes A T2. ☐

Lemma 4.5.7: Comparison with normed spaces
Let A be a set.

Then

- A is a seminormed $\mathbb{K}$-module if and only if A Vs LC with a finite number of seminorms,
- A ‖·‖ $/\mathbb{K}$ if and only if A Vs T2 LC with a finite number of seminorms.

Proof. Suppose A is a seminormed $\mathbb{K}$-module, then there exists a single seminorm $\|\cdot\| : A \to \mathbb{R}$ defining the topology on A as per Definition (4.2.4). However, this is precisely the initial topology of $\|\cdot\|$. By Lemma (4.2.5), A Vs $/\mathbb{K}$ and since the topology on A is the intial topology of $\|\cdot\|$, A LC by Lemma (4.5.3). Therefore A Vs $/\mathbb{K}$ LC.

Suppose conversely that A Vs LC with a finite number of seminorms $\{\|\cdot\|_1, \ldots, \|\cdot\|_k\}$. Then $\|\cdot\| : A \to \mathbb{R}$ defined by

$$\|a\| := \|a\|_1 + \ldots + \|a\|_k$$

is a seminorm on A, the initial topology of which coincides with that of the finite number of seminorms. Hence A is a seminormed $\mathbb{K}$-module.

By Lemma (4.5.6), A ‖·‖ if and only if $\|\cdot\|$ is a norm if and only if A T2. ☐

Lemma 4.5.8
Let A Vs $/\mathbb{K}$ LC.

Then A is pseudometrisable if and only if the collection of open abc neighbourhoods giving rise to local convexity is countable.

Furthermore, if this is the case, then the pseudometric may be assumed to satisfy $d(a_1 + a_3, a_2 + a_3) = d(a_1, a_2)$ for all $a_1, a_2, a_3 \in A$, and A d(·,·) if and only if A T2.

Proof. Suppose A admits a pseudometric d which generates A's topology. Let $\mathcal{A}$ be the basis of open neighbourhoods of 0 arising from local convexity. Then for all $k \in \mathbb{N}$, $k \geq 1$, the set $B_A(0, \frac{1}{k})$ is an open neighbourhood of 0, so for each $k \geq 1$ there exists a $U_k \in \mathcal{A}$ with $0 \in U_k \subseteq B_A(0, \frac{1}{k})$. The countable collection $\mathcal{A}_1 := \{U_1, U_2, \ldots\}$ is by this construction again a basis of open neighbourhoods of 0 that gives rise to local convexity, by Lemma (4.1.6) generates A's topology, and by Lemma (4.5.3) corresponds to a countable collection of seminorms.

Suppose A is locally convex because of a countable collection of seminorms $\|\cdot\|_1, \|\cdot\|_2, \ldots$. Define the function $d : A \times A \to \mathbb{R}$ by

$$d(a_1, a_2) := \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|a_2 - a_1\|_k}{1 + \|a_2 - a_1\|_k}. \tag{4.2}$$

Then for all $a_1, a_2 \in A$, $d(a_1, a_2) \leq \sum_{k=1}^{\infty} \frac{1}{2^k} 1 = 1$, so $d(a_1, a_2) \in \mathbb{R}$. Clearly $d(a_1, a_2) \geq 0$, $d(a_1, a_2) = d(a_2, a_1)$ and $d(a_1, a_1) = \sum_{k=1}^{\infty} 0 = 0$. Because $\|a_3 - a_1\|_k = \|(a_3 - a_2) + (a_2 - a_1)\|_k \leq \|a_3 - a_2\|_k + \|a_2 - a_1\|_k$ and $x/(1+x) + y/(1+y) \geq (x+y)/(1+x+y)$ for $x, y \in [0, \infty[$ we also find $d(a_1, a_3) \leq d(a_1, a_2) + d(a_2, a_3)$. So d is a pseudometric on A. Now note that for any $\epsilon \in]0, \infty[$, $a \in B_A(0, \epsilon)$ if and only if $d(0, a) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|a\|_k}{1 + \|a\|_k} < \epsilon$. There exists an $l \geq 1$ such that $2^{-l} < \epsilon/2$, so as $\sum_{k=l}^{\infty} \frac{1}{2^k} = \sum_{k=0}^{\infty} \frac{1}{2^{k+l}} = 2^{1-l}$ we see that if $\|a\|_1, \ldots, \|a\|_l < \epsilon/2$, then $\sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|a\|_k}{1+\|a\|_k} < \sum_{k=1}^{l} \frac{1}{2^k} \epsilon/2 + \sum_{k=l+1}^{\infty} \frac{1}{2^k} < \epsilon/2 + 2^{1-(l+1)} < \epsilon$, so $\|\cdot\|_1^{-1}(] - 1, \epsilon/2[) \cap \ldots \cap \|\cdot\|_l^{-1}(] - 1, \epsilon/2[) \subseteq B_A(0, \epsilon)$. Conversely, for any $\epsilon \in]0, \infty[$ and $l \in \mathbb{N}$, $l \geq 1$ we can choose $\delta = \frac{\min\{1, \epsilon\}}{2^{l+2}} \in]0, \infty[$ to obtain that if $d_A(0, a) < \delta$, then $\|a\|_k < \epsilon$ for all $k \leq l$. For if $\|a\|_k \geq \epsilon$ for a certain $k \leq l$, then $d_A(0, a) \geq \frac{1}{2^k} \frac{\|a\|_k}{1+\|a\|_k} \geq \frac{1}{2^k} \frac{\epsilon}{1+\epsilon} \geq \frac{1}{2^{k+1}} \min\{1, \epsilon\} > \delta$, contradiction. Therefore $B_A(0, \delta) \subseteq \|\cdot\|_1^{-1}(] - 1, \epsilon[) \cap \ldots \cap \|\cdot\|_l^{-1}(] - 1, \epsilon[)$. Because of this and Lemma (4.1.6), the topology generated by d coincides with the initial topology of the seminorms and hence A is pseudometrisable.

From Equation (4.2) it is clear that $d(a_1 + a_3, a_2 + a_3) = d(a_1, a_2)$ for all $a_1, a_2, a_3 \in A$.

Furthermore, we see from Equation (4.2) that $d(a_1, a_2) = 0$ if and only if for all $k \geq 1$ we have $\|a_2 - a_1\|_k = 0$. Therefore d is a metric if $\forall a \in A : (\forall k \geq 1 : \|a\|_k = 0) \to a = 0$, which by Lemma (4.5.6) is equivalent to A being **T2**. Conversely if d is a metric, then A **T2** by Theorem (2.5.10). ☐

Just as with metric spaces we can talk about completeness for locally convex topological vector spaces, where we demand completeness with respect to all seminorms.

Definition 4.5.9: Uniform completeness (UC)
Let A **Vs** $/\mathbb{K}$ **LC**.

Then we call A *uniformly complete* (denoted by A **UC**) if A is complete with respect to each seminorm, that is if for all sequences $x : \mathbb{N} \to A$ we have that x is convergent in A if

$$\forall i \in I : \forall \epsilon \in]0, \infty[: \exists k \in \mathbb{N} : \forall l, m \geq k : \|x_l - x_m\|_i < \epsilon.$$

This notion is compatible with our original definition (Definition (2.5.14)) of completeness as is shown in the following lemma.

Lemma 4.5.10

Let A TVS $/\mathbb{K}$ LC T2.

If A UC and the collection of abc neighbourhoods giving rise to local convexity is countable, then A FS $/\mathbb{K}$.

Proof. Suppose A UC and has a countable collection of abc neighbourhoods. Then as A T2, A is metrisable by Lemma (4.5.8), denote the metric by d and assume the metric satisfies $d(a_1 + a_3, a_2 + a_3) = d(a_1, a_2)$ for all $a_1, a_2, a_3 \in A$ (translation invariance).

Let $x : \mathbb{N} \to A$ be any Cauchy sequence with respect to the metric d. Let $i \in I$ and $\epsilon \in]0, \infty[$ be arbitrary. As $\|\cdot\|_i : A \to \mathbb{R}$ C, there exists a $\delta \in]0, \infty[$ such that $B_A(0, \delta) \subseteq \| \cdot \|_i^{-1}(] - 1, \epsilon[)$. Because x is Cauchy there exists a $k \in \mathbb{N}$ such that for all $l, m \geq k$ we have $d(x_l, x_m) < \delta$. But as $d(x_l - x_m, 0) = d(x_l, x_m) < \delta$, $x_l - x_m \in B_A(0, \delta)$, so $\|x_l - x_m\|_i \in] - 1, \epsilon[$ and hence $\|x_l - x_m\|_i < \epsilon$. A UC by assumption and the above is true for all $i \in I$ and $\epsilon \in]0, \infty[$, so x is convergent. Because this is true for all Cauchy sequences in A, A is complete. Therefore A FS. □

Example 4.5.11: $\mathbb{K}^k$ TVS $/\mathbb{K}$ T2 LC UC

Let $k \in \mathbb{N}$.

Then the set $\mathbb{K}^k$, together with the norm $\|(x_1, \ldots, x_k)\| := \sqrt{\sum_{l=1}^{k} |x_l|^2}$ is TVS $/\mathbb{K}$ T2 LC (with just a single seminorm). Because $\mathbb{R}$ and $\mathbb{C}$ are complete, $\mathbb{K}^k$ UC.

$\mathbb{Q}^k$ is also TVS $/\mathbb{Q}$ T2 LC, but not UC as $\mathbb{Q}$ is not complete.

Lemma 4.5.12

Let A TVS $/\mathbb{K}$, $B \leq A$.

If A LC, then A/B LC. Furthermore, A/B T2 if and only if $B \subseteq A$ is closed.

Proof. Suppose A LC due to a collection of seminorms $\{\| \cdot \|_i | i \in I\}$. Define $\| \cdot \|_i' : A/B \to \mathbb{R}$ for all $i \in I$ by

$$\|[a]\|_i' := \inf\{\|a_1\|_i \in \mathbb{R} | a_1 \in [a]\}.$$

Note that $0 \leq \|[a]\|_i' \leq \|a\|_i < \infty$, so $\|[a]\|_i' \in [0, \infty[$ for all $[a] \in A/B$.

Let $\alpha \in \mathbb{K}$ and $[a] \in A/B$. For $\alpha \neq 0$, $a_1 \in [a]$ if and only if $a_1 = a + b$ for $b \in B$ if and only if $\alpha a_1 = \alpha\, a + \alpha\, b$ for $b \in B$ if and only if $\alpha\, a_1 \in [\alpha\, a]$. Hence, as $\|\alpha\, a_1\|_i = |\alpha|\, \|a_1\|_i$, we find $\|[\alpha\, a]\|_i' = |\alpha|\, \|[a]\|_i'$. Otherwise, if $\alpha = 0$, then $[\alpha\, a] = [0]$ and as $\|0\|_i = 0$, $0 \in B = [0]$, we see that $\|[0\, a]\|_i' = \|[0]\|_i' = 0 = |0| \|[a]\|_i'$. Hence $\|\alpha[a]\|_i' = |\alpha| \|[a]\|_i'$ for all $\alpha \in \mathbb{K}$, $[a] \in A/B$.

Let $[a_1], [a_2] \in A/B$. Then for all $a_3 \in [a_1], a_4 \in [a_2]$ there exist $b_3, b_4 \in B$ such that $a_3 = a_1 + b_3$, $a_4 = a_2 + b_4$, so $\|a_3\|_i + \|a_4\|_i = \|a_1 + b_3\|_i + \|a_2 + b_4\|_i \geq \|a_1 + a_2 + (b_3 + b_4)\|_i$. As $b_3 + b_4 \in B$, $a_1 + a_2 + (b_3 + b_4) \in [a_1 + a_2]$. So for any $a_3 \in [a_1], a_4 \in [a_2]$ there exists an $a_5 \in [a_1 + a_2]$ such that $\|a_5\|_i \leq \|a_3\|_i + \|a_4\|_i$. Hence $\|[a_1] + [a_2]\|_i' = \|[a_1 + a_2]\|_i' \leq \|[a_1]\|_i' + \|[a_2]\|_i'$ for all $[a_1], [a_2] \in A/B$.

Therefore $\| \cdot \|_i'$ is a seminorm for all $i \in I$.

Let $i \in I$. By definition $\|[a]\|_i' = \inf\{\|a - b\|_i \in \mathbb{R} | b \in B\}$ and hence (in exactly the same fashion as for Lemma (2.5.9)) the map $A \to \mathbb{R} : a \mapsto \|[a]\|_i'$ C. Therefore, as A/B is equipped with the final topology of $a \mapsto [a]$, $\| \cdot \|_i'$ C. Therefore the topology of A/B is as least as large as the initial topology of all the $\| \cdot \|_i'$, $i \in I$.

Let V be an open neighbourhood of $[0]$ in A/B. Then $(a \mapsto [a])^{-1}(V) \subseteq A$ is an open neighbourhood of 0 in A. Since A LC (use Lemma (2.1.19)), there exist $i_1, \ldots, i_k \in I$ and an $\epsilon \in]0, \infty[$ such that $0 \in \|\cdot\|_{i_1}^{-1}(]-1,\epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}^{-1}(]-1,\epsilon[) \subseteq (a \mapsto [a])^{-1}(V)$. As $a \mapsto [a]$ is an open map by Lemma (4.1.9), we find that therefore

$$[0] \in V_1 := (a \mapsto [a])(\|\cdot\|_{i_1}^{-1}(]-1,\epsilon[) \cap \ldots \|\cdot\|_{i_k}^{-1}(]-1,\epsilon[)) \subseteq V$$

and V_1 is an open neighbourhood of $[0]$ in A/B. Now $a \in \|\cdot\|_i^{-1}(]-1,\epsilon[) + B$ iff for some $b \in B$, $\|a+b\|_i < \epsilon$ iff $\inf\{\|a+b\|_i \in \mathbb{R} | b \in B\} < \epsilon$ iff $[a] \in \|\cdot\|_i'^{-1}(]-1,\epsilon[)$, so (use $(a \mapsto [a])(B) = [0]$)

$$[0] \in \|\cdot\|_{i_1}'^{-1}(]-1,\epsilon[) \cap \ldots \cap \|\cdot\|_{i_k}'^{-1}(]-1,\epsilon[) = V_1 \subseteq V.$$

Hence the initial topology generated by the $\|\cdot\|_i'$, $i \in I$ is at least as large as the topology of A/B.

Therefore the topology of A/B equals the initial topology of the $\|\cdot\|_i'$, $i \in I$ and hence A/B LC.

By Lemma (4.1.9) A/B T2 if and only if $B \subseteq A$ is closed. $\square$

Lemma 4.5.13

Let A TVS $/\mathbb{K}$.

Then A' T2 LC.

Proof. Define $\|\cdot\|_a' : A' \to \mathbb{R}$ for all $a \in A$ by

$$\|f\|_a' := |f(a)|.$$

Then clearly $\|f\|_a' = |f(a)| \geq 0$, $\|\alpha f\|_a' = |(\alpha f)(a)| = |\alpha f(a)| = |\alpha||f(a)| = |\alpha|\|f\|_a'$, $\|f+g\|_a' = |(f+g)(a)| = |f(a)+g(a)| \leq |f(a)| + |g(a)| = \|f\|_a' + \|g\|_a'$, so all $\|\cdot\|_a'$ are seminorms on A'. Since the evaluation $f \mapsto f(a) : A' \to \mathbb{K}$ C for all $a \in A$ and $|\cdot| : \mathbb{K} \to \mathbb{R}$ C, $\|\cdot\|_a'$ C for all $a \in A'$. Because of this the topology on A' is at least as large as the initial topology generated by the seminorms $\|\cdot\|_a'$, $a \in A$.

On the other hand, let U' be an open neighbourhood of 0 in A'. Then by Lemma (2.1.19) and definition of the topology of A', there exist $a_1, \ldots, a_k \in A$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, $\beta_1, \ldots, \beta_k \in]0, \infty[$ such that $0 \in (f \mapsto f(a_1))^{-1}(B_{\mathbb{K}}(\alpha_1, \beta_1)) \cap \ldots \cap (f \mapsto f(a_k))^{-1}(B_{\mathbb{K}}(\alpha_k, \beta_k)) \subseteq U'$. Let $1 \leq l \leq k$, as $0 \in (f \mapsto f(a_l))^{-1}(B_{\mathbb{K}}(\alpha_l, \beta_l))$ we have that $|0(a_l) - \alpha_l| = |\alpha_l| < \beta_l$. Choose $\gamma_l := \beta_l - |\alpha_l| \in]0, \infty[$, then if $f \in \|\cdot\|_{a_l}'^{-1}(]-1, \gamma_l[)$, $|f(a_l)| < \gamma_l$, so $|f(a_l) - \alpha_l| \leq |f(a_l)| + |\alpha_l| < \beta_l - |\alpha_l| + |\alpha_l| = \beta_l$, hence $f \in (f \mapsto f(a_l))^{-1}(B_{\mathbb{K}}(\alpha_l, \beta_l))$. Therefore $0 \in \|\cdot\|_{a_1}'^{-1}(]-1,\gamma_1[) \cap \ldots \cap \|\cdot\|_{a_k}'^{-1}(]-1,\gamma_k[) \subseteq (f \mapsto f(a_1))^{-1}(B_{\mathbb{K}}(\alpha_1, \beta_1)) \cap \ldots \cap (f \mapsto f(a_k))^{-1}(B_{\mathbb{K}}(\alpha_k, \beta_k)) \subseteq U'$. Hence the topology of A' is at most as large as the initial topology generated by the seminorms $\|\cdot\|_a'$, $a \in A$. So the seminorms $\|\cdot\|_a'$, $a \in A$ generate the topology of A'. Therefore A' LC.

Let $f \in A'$ and suppose that for all $a \in A$, $\|f\|_a' = |f(a)| = 0$. Then for all $a \in A$, $f(a) = 0$, so $f = 0$. Therefore, by Lemma (4.5.6), A' T2. $\square$

The seminorms from Lemma (4.5.2) permit us to rephrase Theorem (4.2.6) into a very convenient form for locally convex topological vector spaces. This form, Theorem (4.5.14), permits us to translate analysis on our topological vector space to analysis on $\mathbb{K}$ by applying elements of the topological dual to the points we are investigating.

Theorem 4.5.14: Hahn-Banach

Let A [TVS] $/\mathbb{K}$ [T2] [LC].

Then for any $a_1, a_2 \in A$, we have $a_1 = a_2$ if and only if for all $f \in A'$, $f(a_1) = f(a_2)$.

Proof. Because any $f \in A'$ is a function by definition, clearly $f(a_1) = f(a_2)$ if $a_1 = a_2$. Now suppose $a_1 \neq a_2$. As all $f \in A'$ are (l) we may by translating by $-a_1$ suppose that $a_1 = 0$ and $a_2 = a$ for $a \neq 0$. Because $a \neq 0$ and A [T2], by Lemma (4.5.6) there exists an $i \in I$ such that $\|a\|_i > 0$. Consider the map

$$B := \{\alpha\, a \mid \alpha \in \mathbb{K}\} \leq A \qquad g : B \to \mathbb{K} : \alpha\, a \mapsto \alpha\, \|a\|_i.$$

Then g (l) (c) and for all $\alpha\, a \in B$ we have $|g(\alpha\, a)| = \|a\|_i\, |\alpha| = \|\alpha\, a\|_i$. By Theorem (4.2.6) ($B \leq A$, $g \in B^*$, $g \leq \|\cdot\|_i|_B$, $\|\cdot\|_i$ seminorm) there exists an $h \in A^*$ satisfying $h|_B = g$ and $|h(a_1)| \leq \|a_1\|_i$ for all $a_1 \in A$.

Because of this ($\|\cdot\|_i$ (c) by the initial topology of the seminorms on A), $0 \leq \lim_{a_1 \to 0} |h(a_1)| \leq \lim_{a_1 \to 0} \|a_1\|_i = \|0\|_i = 0$, so $\lim_{a_1 \to 0} h(a_1) = 0$ and therefore, as h (l), h (c). Now $h(0) = 0$ and $h(a) = g(a) = \|a\|_i > 0$, so $h(0) \neq h(a)$.

Therefore if $a_1 \neq a_2$, there exists an $h \in A'$ such that $h(a_1) \neq h(a_2)$. □

Because of this theorem, we can also give a stronger, continuous variant of Lemma (3.3.20).

Theorem 4.5.15: Dual of the dual

Suppose $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$, and let A [TVS] $/\mathbb{K}$ [T2] [LC].

Then the map

$$f : A \to (A')' : a \mapsto (g \mapsto g(a))$$

is (c) (l) $/\mathbb{K}$ and bijective.

Proof. We follow [Bou1955]. By Lemma (3.3.20) we already know that f (l) $/\mathbb{K}$.

For continuity note that for any $g \in A'$ the composition of $(h \mapsto h(g)) \circ f$ is given by $((A'' \to \mathbb{K} : h \mapsto h(g)) \circ f)(a) = f(a)(g) = g(a)$, so $(h \mapsto h(g)) \circ f = g$ (c) for any $g \in A'$. By definition of the initial topology on $(A')'$ this makes $f : A \to (A')'$ (c).

Let $a_1, a_2 \in A$ and suppose $f(a_1) = f(a_2)$. Then for any $g \in A'$ we have $g(a_1) = f(a_1)(g) = f(a_2)(g) = g(a_2)$, hence by Theorem (4.5.14) $a_1 = a_2$, so f is injective.

Let $g \in (A')'$, then $g : A' \to \mathbb{K}$ (c) (l) $/\mathbb{K}$. Hence by Lemma (4.5.5) and Lemma (4.5.13) there exists an $\alpha \in]0, \infty[$ and $a_1, \ldots, a_k \in A$ such that for any $h \in A'$ we have $|g(h)| \leq \alpha\,(\|h\|_{a_1} + \ldots + \|h\|_{a_k}) = \alpha\,(|h(a_1)| + \ldots + |h(a_k)|)$. Now let $g_l : A' \to \mathbb{K} : h \mapsto h(a_l)$ for $1 \leq l \leq k$, then $g_1, \ldots, g_k \in (A')'$ and for any $h \in A'$ we have $|g(h)| \leq \alpha\,(|g_1(f)| + \ldots + |g_k(f)|)$. Suppose $h \in \ker g_1 \cap \ldots \cap \ker g_k$, then $|g(h)| \leq \alpha\,(0 + \ldots + 0) = 0$, so $g(h) = 0$ and therefore $h \in \ker g$. Hence $\ker g_1 \cap \ldots \cap \ker g_k \leq \ker g$. Now using the notation of Theorem (3.3.19) for A^* and $(A^*)^*$, this implies that $\langle g \rangle_{\mathbb{K}}^{\perp} = \ker g \geq \ker g_1 \cap \ldots \cap \ker g_k = \langle g_1, \ldots, g_k \rangle_{\mathbb{K}}^{\perp}$. Since these are finite dimensional we therefore have $\langle g \rangle_{\mathbb{K}} = (\langle g \rangle_{\mathbb{K}}^{\perp})^{\perp} \leq (\langle g_1, \ldots, g_k \rangle_{\mathbb{K}}^{\perp})^{\perp} = \langle g_1, \ldots, g_k \rangle_{\mathbb{K}}$. Hence there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ such that $g = \alpha_1\, g_1 + \ldots + \alpha_k\, g_k$, so for $h \in A'$ we have $g(h) = \alpha_1\, g_1(h) + \ldots + \alpha_k\, g_k(h) = \alpha_1\, h(a_1) + \ldots + \alpha_k\, h(a_k) = h(\alpha_1\, a_1 + \ldots + \alpha_k\, a_k)$. Therefore, let $a := \alpha_1\, a_1 + \ldots + \alpha_k\, a_k \in A$, then $g(h) = h(a) = f(a)(h)$ for all $h \in A'$, so $f(a) = g$. So for any $g \in (A')'$ there exists an $a \in A$ such that $f(a) = g$ and this makes f surjective. $\square$

As well as a continuous variant of Theorem (3.3.19) for topological vector spaces.

Theorem 4.5.16: Duality

Let A TVS $/\mathbb{K}$ T2 LC.

Denote for any $B \leq A$ the set

$$B^\perp := \{f \in A' \mid \forall b \in B : f(b) = 0\} \leq A',$$

and for any $C \leq A'$ the set

$$C^\perp := \{a \in A \mid \forall f \in C : f(a) = 0\} \leq A.$$

- Then for all $B \leq A$, $C \leq A'$ we have
$$B \leq (B^\perp)^\perp, \qquad C \leq (C^\perp)^\perp.$$
- For all $B_1 \leq B_2 \leq A$, $C_1 \leq C_2 \leq A'$ we have
$$B_1^\perp \geq B_2^\perp, \qquad C_1^\perp \geq C_2^\perp.$$
- For all $B \leq A$, $C \leq A'$ we have
$$(B^\perp)^\perp = \overline{B} \subseteq A, \qquad (C^\perp)^\perp = \overline{C} \subseteq A'.$$

Proof.
- Show in the same way as in Theorem (3.3.19).
- Show in the same way as in Theorem (3.3.19).
- Let $B \leq A$. We already know that $B \leq (B^\perp)^\perp$. Suppose there exists a $b \in \overline{B}$ for which $b \notin (B^\perp)^\perp$. Then there is an $f \in B^\perp$ with $f(b) \neq 0$. Pick $\epsilon := |f(b)|/2 \in]0, \infty[$. As f C (since $f \in A'$), $U := f^{-1}(B_{\mathbb{K}}(f(b), \epsilon)) \subseteq A$ is an open neighbourhood of b in A. Hence ($b \in \overline{B}$, Lemma (2.1.3)) there exists a $b_1 \in B \cap U$. Now as $b_1 \in B$, $f(b_1) = 0$ ($f \in B^\perp$), but on the other hand, $b_1 \in U$, so $f(b_1) \neq 0$, a contradiction is reached: such a b cannot exist and therefore necessarily $\overline{B} \subseteq (B^\perp)^\perp$.

  Let $b \notin \overline{B}$, then $[b] \neq [0] \in A/\overline{B}$. By Lemma (4.5.12) we see that $A/\overline{B}$ T2 LC, therefore by Theorem (4.5.14) there exists a $g \in (A/\overline{B})'$ such that $g([b]) \neq g([0]) = 0$. Define $f : A \to \mathbb{K} : a \mapsto g([b])$, then $f \in A'$ (as g C I) and for any $b_1 \in B$ we have $f(b_1) = g([b_1]) = g([0]) = 0$. Hence $f \in B^\perp$, however $f(b) = g([b]) \neq 0$, so $b \notin (B^\perp)^\perp$.

  Therefore $\overline{B} = (B^\perp)^\perp$.

  Let $C \leq A'$. Let $f \in \overline{C}$ and suppose $f \notin (C^\perp)^\perp$, then there exists an $a \in C^\perp$ such that $f(a) \neq 0$. As the topology on A' is generated by the seminorms $\|\cdot\|_{a_1} : f_1 \mapsto |f_1(a_1)|$ for all $a_1 \in A$ (Lemma (4.5.13)) we have for $\epsilon := |f(a)|/2 \in]0, \infty[$ that $V := \|\cdot\|_a^{-1}(B_{\mathbb{K}}(f(a), \epsilon))$ is an open neighbourhood of f in A'. As $f \in \overline{C}$ (Lemma (2.1.3)), there exists an $f_1 \in C \cap V$. As $f_1 \in C$, $a \in C^\perp$, we have $f_1(a) = 0$, while on the other hand $f_1 \in V$, so $f_1(a) \neq 0$: we reach a contradiction. Therefore such an f cannot exist and $\overline{C} \subseteq (C^{\perp})^{\perp}$.

Let $f \notin \overline{C}$, then $[f] \neq [0] \in A'/\overline{C}$, so by Theorem (4.5.14) we obtain a $g \in (A'/\overline{C})'$ for which $g([f]) \neq g([0]) = 0$. g in turn gives us a map $(A' \to \mathbb{K} : f_1 \mapsto g([f_1])) \in A''$ which by Theorem (4.5.15) corresponds to a unique $a \in A$ such that $f_1(a) = g([f_1])$ for all $f_1 \in A'$. For any $f_1 \in C$ we have that $f_1(a) = g([f_1]) = g([0]) = 0$, so $a \in C^{\perp}$. On the other hand for our f, $f(a) = g([f]) \neq 0$, so $f \notin (C^{\perp})^{\perp}$.

Therefore $\overline{C} = (C^{\perp})^{\perp}$.

$\square$

Chapter 5

Analysis

Now we are ready to introduce the notions of differentiation (approximating a given function near a given point as well as possible by a linear map), and integration which makes it possible (Theorem (5.3.8)) to recover a function from its derivative and to approximate functions by a sum of 1-l, 2-l, ... maps[1] as a generalisation of approximating a function by its l derivative (this is done in Corollary (5.3.12)).

We will also prove important and very useful existence theorems for Ba spaces: Theorem (5.5.8), Theorem (5.5.9), and Theorem (5.5.10).

5.1 Differentiation

Definition 5.1.1: Differentiability (d, $C^1(U,B)$)
Let A, B TVS $/\mathbb{K}$ T2 LC, and $U \subseteq A$ open.

Then we call a map $f : U \to B$ *differentiable at* a (denoted by f d a) for $a \in U$ if there exists a map $D_a f : A \to B$ c l $/\mathbb{K}$ (called the *derivative of* f *at* a), and a map $\epsilon_{f,a} : (U - a) \to B$ such that for all $a_1 \in (U - a)$ we have

$$f(a + a_1) = f(a) + D_a f(a_1) + \epsilon_{f,a}(a_1) \tag{5.1}$$

and for all $a_1 \in A$,

$$\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} = 0.$$

If f d a for all $a \in U$ we write f d U.

If f d U and the map

$$Df : U \times A \to B : (a, a_1) \mapsto D_a f(a_1)$$

is c, we say that f *is continuously differentiable on* U (denoted by $f \in C^1(U,B)$).

We demand local convexity to ensure that all lines $a + \alpha\, a_1$ are contained in U for small enough $\alpha \in \mathbb{K}$ if U is an open neighbourhood of a in A, as well as T2 to ensure uniqueness of the approximation.

[1]The function's Taylor sequence.

Note that Equation (5.1) really is an expression of the fact that the linear map $D_a f$ is the best linear approximation of our function f near the point a. The 'rest term' $\epsilon_{f,a}(a_1)$ contains the part of f that goes to 0 in a 'faster than linear way' as $a_1 \to 0$.

Lemma 5.1.2: Differentiability implies continuity
Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, $f : U \to B$, and $a \in U$.

If f d a, then f c a.

In particular, if $f \in C^1(U, B)$, then f c.

Proof. Let V be any neighbourhood of 0 in B, as

$$\lim_{(\alpha,a_1)\to(0,0)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} = 0,$$

there exists a $\delta \in]0, 1[$ and an open abc neighbourhood U_1 of 0 in A such that for all $\alpha \in B_{\mathbb{K}}(0, \delta)$ and $a_1 \in U_1$ we have $\epsilon_{f,a}(\alpha\, a_1) \in \alpha V$.

Let $a_2 \in \frac{\delta}{2} U_1$, then $a_2 = \frac{\delta}{2} a_1$ for some $a_1 \in U_1$, so $\epsilon_{f,a}(a_2) = \epsilon_{f,a}(\frac{\delta}{2} a_1) \in \frac{\delta}{2} V \subseteq V$ (as $\delta \in]0, 1[$).

Hence, for any open neighbourhood V of 0 in B there exists an open neighbourhood $\frac{\delta}{2} U_1$ of 0 in A such that for all $a_2 \in \frac{\delta}{2} U_1$ we have $\epsilon_{f,a}(a_2) \in V$. Therefore $\lim_{a_1\to 0} \epsilon_{f,a}(a_1) = 0$.

Using this we see that

$$\begin{aligned}
\lim_{a_1\to 0} f(a + a_1) &= \lim_{a_1\to 0} \Big(f(a) + D_a f(a_1) + \epsilon_{f,a}(a_1)\Big) \\
&= f(a) + \lim_{a_1\to 0} D_a f(a_1) + \lim_{a_1\to 0} \epsilon_{f,a}(a_1) \\
&= f(a) + D_a f(0) + 0 \\
&= f(a),
\end{aligned}$$

because all limits exist, addition is continuous, and $D_a f$ c l. Hence f c a. ☐

Example 5.1.3: Differentiability is a local property
Consider $f : \mathbb{R} \to \mathbb{R}$ from Example (2.1.15). Then f d 0 ($\lim_{\alpha\to 0} \frac{1}{\alpha}(f(0 + \alpha) - f(0)) = 0$), but for all $\alpha \in \mathbb{R} \setminus \{0\}$ we have that not f d α.

So this function is differentiable in just a single point.

Lemma 5.1.4: Uniqueness of the derivative
Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, $a \in U$, $f : U \to B$.

If the map $g : A \to B$ given by

$$g(a_1) := \lim_{(\alpha,a_2)\to(0,a_1)} \frac{1}{\alpha}\Big(f(a + \alpha\, a_2) - f(a)\Big)$$

is defined for all $a_1 \in A$ and is c l $/\mathbb{K}$, then f d a and $D_a f = g$.

Furthermore, if f d a, then for all $a_1 \in A$,

$$D_a f(a_1) = \lim_{\alpha\to 0} \frac{1}{\alpha}\Big(f(a + \alpha\, a_1) - f(a)\Big).$$

In particular, the derivative is unique.

Proof. Suppose that for all $a_1 \in A$ we have existence of the limit

$$g(a_1) = \lim_{(\alpha,a_2)\to(0,a_1)} \frac{1}{\alpha}\Big(f(a+\alpha\, a_2) - f(a)\Big)$$

in B, and $g : A \to B$ ⓒ ⓛ $/\mathbb{K}$. From the definition of g we have for any $a_1 \in U - a$ that

$$\lim_{(\alpha,a_2)\to(0,a_1)} \frac{1}{\alpha}\Big(f(a+\alpha a_2) - f(a) - \alpha\, g(a_2)\Big) = 0,$$

so choosing for all $a_1 \in U - a$ the function

$$\epsilon_{f,a}(a_1) := f(a + a_1) - f(a) - g(a_1)$$

we see that

$$f(a + a_1) = f(a) + g(a_1) + \epsilon_{f,a}(a_1)$$

and

$$\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} = 0.$$

Hence f ⓓ a.

Let $D_a f : A \to B$ satisfy Equation (5.1) (existence of at least one such map is guaranteed by the fact that f ⓓ a). Because A LC there exists an open abc neighbourhood U_1 of 0 in A such that $U_1 + a \subseteq U$. Let $a_1 \in A$, then because $\lim_{\alpha\to 0} \alpha\, a_1 = 0$ there exists a $\delta \in]0,\infty[$ such that for all $\alpha \in B_{\mathbb{K}}(0,\delta)$ we have $\alpha\, a_1 \in U_1$ and hence

$$f(a + \alpha\, a_1) = f(a) + D_a f(\alpha\, a_1) + \epsilon_{f,a}(\alpha\, a_1),$$

so for all $\alpha \neq 0$, $|\alpha| < \delta$, (use $D_a f$ ⓛ),

$$\frac{\epsilon_{f,a}(\alpha\, a_1)}{\alpha} = \frac{1}{\alpha}\Big(f(a + \alpha\, a_1) - f(a)\Big) - D_a f(a_1)$$

and as $\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} = 0$ we obtain

$$D_a f(a_1) = \lim_{(\alpha,a_2)\to(0,a_1)} \frac{1}{\alpha}\Big(f(a + \alpha\, a_2) - f(a)\Big).$$

Therefore (A T2, Lemma (2.2.6)) $D_a f(a_1) = g(a_1)$ for all $a_1 \in A$, so $D_a f = g$ and the derivative is unique.

The last statement now follows from

$$\lim_{(\alpha,a_2)\to(0,a_1)} \frac{1}{\alpha}\Big(f(a + \alpha\, a_2) - f(a)\Big) = \lim_{\alpha\to 0} \frac{1}{\alpha}\Big(f(a + \alpha\, a_1) - f(a)\Big),$$

because the limit on the left hand side exists whenever f ⓓ a. $\square$

Because this definition of differentiability is different from the usual Fréchet or Gâteaux derivative (see for example [Ham1982]), we need to check how it compares to these types of differentiation. Clearly the demands of Definition (5.1.1) are stronger than those made of the Gâteaux derivative (which need not even be linear), but how it compares to the Fréchet derivative is not immediately clear.

Theorem 5.1.5: Compatible differentiation
Let A ‖·‖ $/\mathbb{K}$, B TVS $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, $f : U \to B$, and $a \in U$.

If there exists a map $g : A \to B$ c l $/\mathbb{K}$ such that

$$\lim_{a_1 \to 0} \frac{1}{\|a_1\|_A} \Big(f(a + a_1) - f(a) - g(a_1) \Big) = 0,$$

then f d a and $D_a f = g$.

Proof. By Lemma (4.5.7), A TVS $/\mathbb{K}$ T2 LC.

Fix $a_1 \in A$ and let V be an open abc neighbourhood of 0 in B. Then because A ‖·‖ and the above limit, there exists a $\delta_1 \in]0, \infty[$ such that for all $a_2 \in B_A(0, \delta_1)$ we have

$$f(a + a_2) - f(a) - g(a_2) \in \|a_2\|_A \, \frac{1}{1 + \|a_1\|_A} \, V.$$

Choose $\delta := \delta_1/(1+\|a_1\|_A)$ and $U := B_A(a_1, 1)$ which is an open neighbourhood of a_1 in A. Then for all $\alpha \in B_{\mathbb{K}}(0, \delta)$, $\alpha \neq 0$ and $a_2 \in U$ we have $\|a_2\|_A < 1 + \|a_1\|_A$ and therefore $\|\alpha\, a_2\|_A = |\alpha| \, \|a_2\|_A < \delta\,(1 + \|a_1\|_A) = \delta_1$, so (use Lemma (4.3.5) and the fact that V is abc)

$$\begin{aligned} f(a + \alpha\, a_2) - f(a) - g(\alpha\, a_2) &\in \|\alpha\, a_2\|_A \, \frac{1}{1 + \|a_1\|_A} \, V \\ &\subseteq |\alpha| \, (1 + \|a_1\|_A) \, \frac{1}{1 + \|a_1\|_A} \, V \\ &= \alpha \, \frac{|\alpha|}{\alpha} \, V \\ &= \alpha\, V. \end{aligned}$$

So for all $a_1 \in A$ and open abc neighbourhoods V of 0 in B, there exists a $\delta \in]0, \infty[$ and open neighbourhood U of a_1 in A such that for $\alpha \in B_{\mathbb{K}}(0, \delta)$, $\alpha \neq 0$ and $a_2 \in U$ we have

$$\frac{1}{\alpha} \Big(f(a + \alpha\, a_2) - f(a) - g(\alpha\, a_2) \Big) \in V.$$

Hence for all $a_1 \in A$

$$\lim_{(\alpha, a_2) \to (0, a_1)} \frac{1}{\alpha} \Big(f(a + \alpha\, a_2) - f(a) - g(\alpha\, a_2) \Big) = 0.$$

As g c l we therefore find that f d a.

By Lemma (5.1.4) we furthermore see that $D_a f = g$. □

We immediately obtain the following consequence if B is also a normed space, which shows that the derivative defined in Definition (5.1.1) is at least as general as the notion of Fréchet differentiability (see Corollary (5.5.7)).

Corollary 5.1.6: Compatibility on normed spaces
Let A, B ‖·‖ $/\mathbb{K}$, $U \subseteq A$ open, $f : U \to B$.

Let $a \in U$. If there exists a map $g : A \to B$ c l $/\mathbb{K}$ such that

$$\lim_{a_1 \to 0} \frac{\|f(a + a_1) - f(a) - g(a_1)\|_B}{\|a_1\|_A} = 0,$$

then f d a and $D_a f = g$.

For paths (functions from an interval in $\mathbb{R}$ to A), we can do even better.

Corollary 5.1.7: Compatibility for paths
Let A TVs $/\mathbb{K}$ T2 LC and $S \subseteq \mathbb{R}$ an open interval, $f : S \to A$.

Let $\alpha \in S$. The limit

$$f'(\alpha) = \lim_{\beta \to \alpha} \frac{f(\beta) - f(\alpha)}{\beta - \alpha}$$

exists if and only if f d α, in which case

$$D_\alpha f(\beta) = \beta\, f'(\alpha)$$

for all $\beta \in \mathbb{R}$.

In particular, this expression shows that if the limit $f'(\alpha)$ exists for all $\alpha \in S$, then $S \to A : \alpha \mapsto f'(\alpha)$ c if and only if $f \in C^1(S, A)$.

Proof. Note that even though A may be a TVs $/\mathbb{C}$ while $\mathbb{R}$ itself is a TVs $/\mathbb{R}$, this does not give any problems for linearity of the involved maps since $\mathbb{R} \leq \mathbb{C}$, so we may consider A as a TVs $/\mathbb{R}$ by taking the restriction of scalar multiplication to $\mathbb{R} \times A \subseteq \mathbb{C} \times A$.

Let $\alpha \in S$. Suppose that the limit for $f'(\alpha)$ exists, then

$$\begin{aligned} 0 &= f'(\alpha) - f'(\alpha) \\ &= \lim_{\beta \to \alpha} \left(\frac{f(\beta) - f(\alpha)}{\beta - \alpha} - \frac{(\beta - \alpha)\, f'(\alpha)}{\beta - \alpha} \right) \\ &= \lim_{\beta \to 0} \frac{1}{\beta} \Big(f(\alpha + \beta) - f(\alpha) - \beta\, f'(\alpha) \Big), \end{aligned}$$

so using the fact that we can choose balanced neighbourhoods of 0 in A and the expression can only change sign as $1/\beta = \pm 1/|\beta|$, we find

$$\lim_{\beta \to 0} \frac{1}{|\beta|} \Big(f(\alpha + \beta) - f(\alpha) - \beta\, f'(\alpha) \Big) = 0.$$

Hence we can apply Theorem (5.1.5) with the map $\mathbb{R} \to A : \beta \mapsto \beta\, f'(\alpha)$ c l to conclude that f d α.

Suppose conversely that f d α, then for $1 \in \mathbb{R}$ we have

$$\lim_{(\beta,\gamma) \to (0,1)} \frac{1}{\beta} \Big(f(\alpha + \beta\,\gamma) - f(\alpha) - D_\alpha f(\beta\,\gamma) \Big) = 0.$$

Let U be an open abc neighbourhood of 0 in A, then because of the above limit there exists a $\delta \in]0, \infty[$ such that $|\beta| < \delta$ and $|\gamma - 1| < \delta$ imply

$$f(\alpha + \beta\,\gamma) - f(\alpha) - \beta\,\gamma\, D_\alpha f(1) \in \beta\, U.$$

In particular for $\gamma = 1$ and $|\beta| < \delta$, $\beta \neq 0$,

$$\frac{1}{\beta} \Big(f(\alpha + \beta) - f(\alpha) \Big) - D_\alpha f(1) \in U.$$

Hence

$$\lim_{\beta \to 0} \left(\frac{1}{\beta} \Big(f(\alpha + \beta) - f(\alpha) \Big) - D_\alpha f(1) \right) = 0,$$

so
$$D_\alpha f(1) = \lim_{\beta \to \alpha} \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$
Therefore the limit $f'(\alpha)$ exists and is equal to $D_\alpha f(1)$. □

Now that we have established that Definition (5.1.1) is reasonable, we can determine the properties of the derivative of a function in Theorem (5.1.8) (compare with Theorem (2.1.28)).

Theorem 5.1.8: Operations preserving differentiability
Let A, B Vs $/\mathbb{K}$ T2 LC.

Composition: let C Vs $/\mathbb{K}$, $U \subseteq A$ open, $V \subseteq B$ open, $f : U \to B$, $g : V \to C$, $a \in U$, $f(a) \in V$.

If f d a, g d $f(a)$, then $f \circ g$ d a and
$$D_a(g \circ f) = D_{f(a)} g \circ D_a f.$$
In particular if $f \in C^1(U, B)$ and $g \in C^1(V, C)$ with $f(U) \subseteq V$, then $g \circ f \in C^1(U, C)$.

Addition: let $U \subseteq A$ open, $a \in U$, $f : U \to B$, $g : U \to B$.

If f, g d a, then $g + f$ d a and
$$D_a(g + f) = D_a g + D_a f.$$
In particular if $f, g \in C^1(U, B)$, then $f + g \in C^1(U, B)$.

Scaling: let $U \subseteq A$ open, $a \in U$, $f : U \to B$, $\alpha \in \mathbb{K}$.

If f d a, then $\alpha\, f$ d a and
$$D_a(\alpha\, f) = \alpha\, D_a f.$$
In particular for $f \in C^1(U, B)$ and $\alpha \in \mathbb{K}$, we have $\alpha\, f \in C^1(U, B)$.

Glueing: let $U_1, U_2 \subseteq A$ be both open, $f_1 \in C^1(U_1, B)$, $f_2 \in C^1(U_2, B)$.

If $f_1|_{U_1 \cap U_2} = f_2|_{U_1 \cap U_2}$, then there exists a unique $f \in C^1(U, B)$ such that $f|_{U_1} = f_1$ and $f|_{U_2} = f_2$.

Restricting domain: let $U \subseteq A$ open, and $f \in C^1(U, B)$.

Then for any $U_1 \subseteq U$ open, $f|_{U_1} \in C^1(U_1, B)$.

Constantness: let $U \subseteq A$ abc open, $f : U \to B$.

Then there exists a $b \in B$ such that $f(a) = b$ for all $a \in U$ if and only if $f \in C^1(U, B)$ and $D_a f = 0$ for all $a \in U$.

Linearity: let $f : A \to B$ c l $/\mathbb{K}$.

Then $f \in C^1(A, B)$ and $D_a f = f$ for all $a \in A$.

If conversely $f \in C^1(A, B)$ and there exists a $g : A \to B$ c l $/\mathbb{K}$ such that $D_a f = g$ for all $a \in A$, then there exists a $b \in B$ such that $f(a) = g(a) + b$ for all $a \in A$.

Proof. We will cover this item by item.

Composition: let $C \boxed{\text{Vs}} /\mathbb{K}$, $U \subseteq A$ open, $V \subseteq B$ open, $f : U \to B$, $g : V \to C$, $a \in U$, $f(a) \in V$. Suppose $f \textcircled{d}\, a$, $g \textcircled{d}\, f(a)$, then for $a_1 \in U - a$ we have (use $D_{f(a)}g \textcircled{l}$)

$$\begin{aligned}(g \circ f)(a + a_1) &= g(f(a + a_1)) \\ &= g(f(a) + (D_a f(a_1) + \epsilon_{f,a}(a_1))) \\ &= g(f(a)) + D_{f(a)}g(D_a f(a_1) + \epsilon_{f,a}(a_1)) + \epsilon_{g,f(a)}(D_a f(a_1) + \epsilon_{f,a}(a_1)) \\ &= (g \circ f)(a) + (D_{f(a)}g \circ D_a f)(a_1) + \epsilon_{f\circ g,a}(a_1)\end{aligned}$$

where we have defined

$$\epsilon_{f\circ g,a}(a_1) := D_{f(a)}g(\epsilon_{f,a}(a_1)) + \epsilon_{g,f(a)}(D_a f(a_1) + \epsilon_{f,a}(a_1)).$$

Now for any $a_2 \in A$, $\alpha \in \mathbb{K}$, $\alpha \neq 0$ sufficiently small we obtain (use $D_a f$, $D_{f(a)}g \textcircled{l}$ and Equation (5.1))

$$\begin{aligned}\frac{\epsilon_{f\circ g,a}(\alpha\, a_2)}{\alpha} = D_{f(a)}g\left(\frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha}\right) \\ + \frac{1}{\alpha}\epsilon_{g,f(a)}\left(\alpha\left(D_a f(a_2) + \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha}\right)\right).\end{aligned}$$

As $D_a f$, $D_{f(a)}g \textcircled{c}$ and for any $a_1 \in A$, $b_1 \in B$

$$\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} = \lim_{(\alpha,b_2)\to(0,b_1)} \frac{\epsilon_{g,f(a)}(\alpha\, b_2)}{\alpha} = 0,$$

we obtain with Lemma (2.1.10) that for any $a_1 \in A$

$$\begin{aligned}\lim_{(\alpha,a_2)\to(0,a_1)} D_{f(a)}g\left(\frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha}\right) &= \lim_{a_2\to 0} D_{f(a)}g(a_2) \\ &= D_{f(a)}g(0) = 0 \\ \lim_{(\alpha,a_2)\to(0,a_1)} \left(D_a f(a_2) + \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha}\right) &= \lim_{a_2\to a_1} D_a f(a_2) \\ &\quad + \lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha} \\ &= D_a f(a_1) + 0.\end{aligned}$$

Hence for all $a_1 \in A$,

$$\begin{aligned}\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f\circ g,a}(\alpha\, a_2)}{\alpha} &= 0 + \lim_{(\alpha,b_2)\to(0,D_a f(a_1))} \frac{1}{\alpha}\epsilon_{g,f(a)}(\alpha\, b_2) \\ &= 0.\end{aligned}$$

Together with the fact that $D_{f(a)}g \circ D_a f \textcircled{c}\textcircled{l} /\mathbb{K}$ as both $D_{f(a)}g$ and $D_a f$ have these properties, we see that $g \circ f \textcircled{d}\, a$. As the derivative is unique by Lemma (5.1.4), we furthermore obtain $D_a(g \circ f) = D_{f(a)}g \circ D_a f$.

Addition: let $U \subseteq A$ open, $a \in U$, $f : U \to B$, $g : U \to B$. Suppose f, g d a, then in the same way as for composition we find for $a_1 \in U - a$ that

$$\begin{aligned}(g+f)(a+a_1) &= g(a+a_1) + f(a+a_1)\\ &= g(a) + D_a g(a_1) + \epsilon_{g,a}(a_1) + f(a) + D_a f(a_1) + \epsilon_{f,a}(a_1)\\ &= (g+f)(a) + (D_a g + D_a f)(a_1) + \epsilon_{g+f,a}(a_1)\end{aligned}$$

where we defined

$$\epsilon_{g+f,a}(a_1) := \epsilon_{g,a}(a_1) + \epsilon_{f,a}(a_1).$$

Now for any $a_1 \in A$ we have

$$\begin{aligned}\lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{g+f,a}(\alpha\, a_2)}{\alpha} &= \lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{g,a}(\alpha\, a_2)}{\alpha} + \lim_{(\alpha,a_2)\to(0,a_1)} \frac{\epsilon_{f,a}(\alpha\, a_2)}{\alpha}\\ &= 0 + 0 = 0.\end{aligned}$$

Clearly $D_a g + D_a f$ c l $/\mathbb{K}$, so $g + f$ d a and by uniqueness $D_a(g+f) = D_a g + D_a f$.

Scaling: let $U \subseteq A$ open, $a \in U$, $f : U \to B$, $\alpha \in \mathbb{K}$. Suppose f d a, then for any $a_1 \in U - a$

$$\begin{aligned}(\alpha\, f)(a+a_1) &= \alpha\, f(a+a_1)\\ &= \alpha\ (f(a) + D_a f(a_1) + \epsilon_{f,a}(a_1))\\ &= (\alpha\, f)(a) + (\alpha\, D_a f)(a_1) + \epsilon_{\alpha\, f,a}(a_1)\end{aligned}$$

where we defined

$$\epsilon_{\alpha\, f,a}(a_1) := \alpha\, \epsilon_{f,a}(a_1).$$

For any $a_1 \in A$ we have

$$\lim_{(\beta,a_2)\to(0,a_1)} \frac{\epsilon_{\alpha\, f,a}(\beta a_2)}{\beta} = \alpha\, 0 = 0.$$

Clearly $\alpha\, D_a f$ c l $/\mathbb{K}$, so $\alpha\, f$ d a and by uniqueness $D_a(\alpha\, f) = \alpha\, D_a f$.

Glueing, restricting domain: Both follow directly from the fact that the definition of differentiability is only made on an (arbitrarily small) open neighbourhood of a given point and hence local.

Constantness: let $U \subseteq A$ abc open, $f : U \to B$. Suppose $f(a) = b$ for all $a \in U$. Let $a \in U$, $a_1 \in U - a$, then

$$\begin{aligned}f(a+a_1) &= b\\ &= f(a) + 0(a_1) + 0.\end{aligned}$$

Since the zero map is c l, f d a and $D_a f = 0$. Now the map $U \times A \to B : (a, a_1) \mapsto D_a f(a_1) = 0$ is constant and hence c, so $f \in C^1(U, B)$.

Suppose conversely that $f \in C^1(U, B)$ and $D_a f = 0$ for all $a \in U$. Let $a \in U$ and $a_1 \in A$. As U is abc, we can for any $g \in B'$ define for sufficiently small $\delta \in]0, \infty[$

$$h :]-\delta, \delta[\to \mathbb{K} : \alpha \mapsto g(f(a + \alpha\, a_1))$$

which is ⓒ as a composition of ⓒ maps. Since f ⓓ a, we have by Lemma (5.1.4) (use that g ⓒⓛ)

$$\begin{aligned}
\lim_{\alpha\to 0}\frac{h(\alpha)-h(0)}{\alpha-0} &= \lim_{\alpha\to 0}\frac{1}{\alpha}\Big(g(f(a+\alpha\, a_1))-g(f(a))\Big)\\
&= \lim_{\alpha\to 0} g\left(\frac{1}{\alpha}\Big(f(a+\alpha\, a_1)-f(a)\Big)\right)\\
&= g\left(\lim_{\alpha\to 0}\frac{1}{\alpha}\Big(f(a+\alpha\, a_1)-f(a)\Big)\right)\\
&= g\left(D_a f(a_1)\right).
\end{aligned}$$

So the limit $h'(0)=g(D_a f(a_1))=g(0)=0$ exists and is equal to zero. Therefore, for any $g\in A'$, and $a_1,a_2\in U$ the continuous map (well-defined because U is convex)

$$i:[0,1]\to\mathbb{K}:\alpha\mapsto g(f(a_1+\alpha\,(a_2-a_1)))$$

is differentiable on $]0,1[$ with derivative equal to zero (consider for each $\alpha\in]0,1[$, $a=a_1+\alpha\,(a_2-a_1)$ and direction a_2-a_1, now take the derivative of $\beta\mapsto g(f(a+\beta\,(a_2-a_1)))$ as above). In the case that $\mathbb{K}=\mathbb{R}$ we have ($i:[0,1]\to\mathbb{R}$ continuous, $i|_{]0,1[}$ differentiable) by the mean value theorem that there exists an $\alpha\in]0,1[$ such that $i(1)-i(0)=i'(\alpha)(1-0)=0$. Hence $i(0)=i(1)$. In the case that $\mathbb{K}=\mathbb{C}$, apply the mean value theorem to $\operatorname{Re} i$ and $\operatorname{Im} i$ separately. This gives $g(f(a_1))=i(0)=i(1)=g(f(a_2))$. Since this is true for all $g\in A'$ and $a_1,a_2\in U$, Theorem (4.5.14) implies that $f(a_1)=f(a_2)$ for all $a_1,a_2\in U$: f is constant.

Linearity: let $f:A\to B$ ⓒⓛ $/\mathbb{K}$. Then for any $a\in A$, $a_1\in A-a=A$ we have

$$\begin{aligned}
f(a+a_1) &= f(a)+f(a_1)\\
&= f(a)+f(a_1)+0.
\end{aligned}$$

As f ⓒⓛ $/\mathbb{K}$, f ⓓ a and by uniqueness $D_a f=f$. Furthermore, the map $A\times A\to B:(a,a_1)\mapsto D_a f(a_1)=f(a_1)$ ⓒ, so $f\in C^1(A,B)$.

Suppose that $f\in C^1(A,B)$ and let $g:A\to B$ ⓒⓛ such that $D_a f=g$ for all $a\in A$. Then the map $h:A\to B:a\mapsto f(a)-g(a)$ satisfies $D_a h=D_a f-g=g-g=0$ for all $a\in A$, so there exists a $b\in B$ such that $h(a)=b$ for all $a\in A$. Hence $f(a)=h(a)+g(a)=g(a)+b$ for all $a\in A$.

□

Definition 5.1.9: Higher order derivatives ($C^k(U,B)$)
Let A, B [Vs] $/\mathbb{K}$ [T2] [LC], $U\subseteq A$ open, and $a\in U$.

Let $k\in\mathbb{N}_0$.

Suppose $k=0$, then f ⓒ U is denoted by $f\in C^0(U,B)$ and if this is the case we write $D^0 f:U\to B:a\mapsto D^0_a f:=f(a)$.

Suppose $k=1$, then we say f is 1 *time differentiable at* a if f ⓓ a. If $f\in C^1(U,B)$, we denote $D^1 f:U\times A\to B:(a,a_1)\mapsto D^1_a f(a_1):=D_a f(a_1)$.

Suppose $k \geq 2$, then we inductively define f to be k *times differentiable at* a if f is $k-1$ times differentiable at a and there exists a map $D_a^k f : A^k \to B$ c k-l /$\mathbb{K}$ such that for all $a_1, \ldots, a_{k-1} \in A$ the map

$$U \to B : a' \mapsto D_{a'}^{k-1} f(a_1, \ldots, a_{k-1})$$

is d a with derivative $A \to B : a_k \mapsto D_a^k f(a_1, \ldots, a_{k-1}, a_k)$.

Suppose $k \geq 2$, then we inductively define f to be k *times continuously differentiable on* U (denoted by $f \in C^k(U, B)$) if $f \in C^{k-1}(U, B)$, and the map

$$D^k f : U \times A^k \to B : (a, a_1, \ldots, a_k) \mapsto D_a^k f(a_1, \ldots, a_k)$$

is c.

We say f is *smooth on* U (denoted by $f \in C^\infty(U, B)$) if $f \in C^k(U, B)$ for all $k \in \mathbb{N}$.

We say f is *analytic on* U (denoted by $f \in C^\omega(U, B)$) if $f \in C^\infty(U, B)$ and for each $a \in U$ there exists an open neighbourhood U_1 of a in U such that for all $a_1 \in U_1 - a$ we have

$$f(a + a_1) = \sum_{k=0}^{\infty} \frac{1}{k!} D_a^k f(\underbrace{a_1, \ldots, a_1}_{k}).$$

Note that this definition enables us to apply Theorem (5.1.8) to k times differentiable functions as well.

The definition of analyticity is motivated by Theorem (5.3.11): a function is analytic if its Taylor series locally gives a complete description of the function.

Example 5.1.10: $C^k(A, B)$ Vs /$\mathbb{K}$

Let A, B Vs /$\mathbb{K}$ T2 LC, $U \subseteq A$ open, $k \in \mathbb{N}$.

Since addition and scalar multiplication on A and B are c we have by Theorem (2.1.28) and Theorem (5.1.8) that for any $f, g \in C^k(U, B)$, $\alpha \in \mathbb{K}$, $f + \alpha\, g \in C^k(U, B)$. This makes $C^k(U, B)$ Vs /$\mathbb{K}$.

Furthermore, from Definition (5.1.9) we know that $C^0(U, B) \supseteq C^1(U, B) \supseteq \ldots$, therefore, as Vs /$\mathbb{K}$ we have

$$C^0(U, B) \geq C^1(U, B) \geq \ldots \geq C^k(U, B) \geq \ldots.$$

Lemma 5.1.11

Let $k \in \mathbb{N}$, A, B_1, ..., B_k Vs /$\mathbb{K}$ T2 LC, $U \subseteq A$ open.

Let $f : U \to B_1 \times \ldots \times B_k : a \mapsto (f_1(a), \ldots, f_k(a))$. Then for $a \in U$, f d a if and only if for all $1 \leq l \leq k$ the map $f_l : U \to B_l$ d a.

If this is the case, then for all $a_1 \in A$

$$D_a f(a_1) = (D_a f_1(a_1), \ldots, D_a f_k(a_1)).$$

In particular, $f \in C^l(U, B_1 \times \ldots \times B_k)$ if and only if for all $1 \leq m \leq l$, $f_m \in C^l(U, B_m)$.

Proof. Suppose f d a and let $1 \leq l \leq k$. Write $g_l : B_1 \times \ldots \times B_k \to B_l : (b_1, \ldots, b_k) \mapsto b_l$, then $g_l \in C^1(B_1 \times \ldots \times B_k, B_l)$ since g_l c l /$\mathbb{K}$ (use Theorem (5.1.8)). Hence g_l d $f(a)$ and therefore by Theorem (5.1.8) $f_l = g_l \circ f$ d a.

Suppose conversely that for all $1 \leq l \leq k$, f_l ⓓ a. Then for $a_2 \in A$ and $\alpha \in \mathbb{K}$ small enough but nonzero we have

$$\frac{1}{\alpha}\Big(f(a+\alpha\, a_2)-f(a)\Big) = \Big(\frac{1}{\alpha}\Big(f_1(a+\alpha\, a_2)-f_1(a)\Big), \dots, \frac{1}{\alpha}\Big(f_k(a+\alpha\, a_2)-f_k(a)\Big)\Big)$$

so letting $(\alpha, a_2) \to (0, a_1)$ we see that as all f_l ⓓ a, we have by Lemma (5.1.4) that f ⓓ a with

$$D_a f(a_1) = (D_a f_1(a_1), \dots, D_a f_k(a_1)).$$

□

Definition 5.1.12: Partial derivative
Let $k \in \mathbb{N}$, $A_1, \dots, A_k$, B Vs$/\mathbb{K}$ T2 LC, $U_1 \subseteq A_1, \dots, U_k \subseteq A_k$ open.

Define $U := U_1 \times \ldots \times U_k$ and let $f \in C^1(U, B)$.

Then we define for $1 \leq l \leq k$, $(a_1, \dots, a_k) \in U$ and $a' \in A$ the *l-th partial derivative of f in direction of a' at $(a_1, \dots, a_k)$* as

$$\frac{\partial}{\partial a_l}\Big(f(a_1, \dots, a_k)\Big)(a') := \Big(\frac{\partial f(a_1, \dots, a_k)}{\partial a_l}\Big)(a') := D_{a_l} g(a')$$

where $g : U_l \to B$ is given by $g(a) := f(a_1, \dots, a_{l-1}, a, a_{l+1}, \dots, a_k)$.

In the case that for certain $1 \leq l \leq k$, $A_l = \mathbb{K}$ we abbreviate

$$\frac{\partial}{\partial a_l}\Big(f(a_1, \dots, a_k)\Big) := \frac{\partial f(a_1, \dots, a_k)}{\partial a_l} := D_{a_l} g(1).$$

For higher order derivatives we define for $m > 1$, $f \in C^m(U, B)$, $i_1, \dots, i_m \in \{1, \dots, k\}$, and $a'_1 \in A_{i_1}, \dots, a'_m \in A_{i_m}$ that

$$\Big(\frac{\partial^m f(a_1, \dots, a_k)}{\partial a_{i_m} \ldots \partial a_{i_1}}\Big)(a'_1, \dots, a'_m) := \frac{\partial}{\partial a_{i_m}}\Big(\Big(\frac{\partial^{m-1} f(a_1, \dots, a_k)}{\partial a_{i_{m-1}} \ldots \partial a_{i_1}}\Big)(a'_1, \dots, a'_{m-1})\Big)(a'_m).$$

This expression is symmetric in permutations $i_{\pi(1)}, \dots, i_{\pi(m)}$, $a'_{\pi(1)}, \dots, a'_{\pi(m)}$ for $\pi \in S^m$ by Theorem (5.1.16).

Lemma 5.1.13: Sum rule
Let $k \in \mathbb{N}$, $A_1, \dots, A_k$, B Vs$/\mathbb{K}$ T2 LC, $U_1 \subseteq A_1, \dots, U_k \subseteq A_k$ open.

Define $U := U_1 \times \ldots \times U_k$ and let $f \in C^1(U, B)$.

Then the map $U \times A_l \to B$ given by

$$(a_1, \dots, a_k, a'_l) \mapsto \Big(\frac{\partial f(a_1, \dots, a_k)}{\partial a_l}\Big)(a'_l)$$

is ⓒ.

Furthermore, for any $(a_1, \dots, a_k) \in U$, $a'_1 \in A_1, \dots, a'_k \in A_k$ we have

$$D_{(a_1, \dots, a_k)} f(a'_1, \dots, a'_k) = \sum_{l=1}^{k} \frac{\partial f(a_1, \dots, a_k)}{\partial a_l}(a'_l). \tag{5.2}$$

Proof. Let us for each $(a_1, \ldots, a_k) \in U$, $1 \leq l \leq k$ use the notation

$$(\widehat{a_l}) := (a_1, \ldots, a_{l-1}, a_{l+1}, \ldots, a_k)$$

and define the map $g_l^{(\widehat{a_l})} : U_l \to B$ by

$$g_l^{(\widehat{a_l})}(a) := f(a_1, \ldots, a_{l-1}, a, a_{l+1}, \ldots, a_k).$$

Suppose $f \in C^1(U, B)$, let $1 \leq l \leq k$, $(a_1, \ldots, a_k) \in U$, and $a_l'' \in A_l$. Then for $\alpha \in \mathbb{K}$ small enough but nonzero

$$\begin{aligned} &\frac{1}{\alpha}\Big(g_l^{(\widehat{a_l})}(a_l + \alpha\, a_l'') - g_l^{(\widehat{a_l})}(a_l)\Big) \\ &\qquad\qquad = \frac{1}{\alpha}\Big(f(a_1, \ldots, a_l + \alpha\, a_l'', \ldots, a_k) - f(a_1, \ldots, a_l, \ldots, a_k)\Big) \end{aligned}$$

So taking for $a_l' \in A_l$ the limit $(\alpha, a_l'') \to (0, a_l')$ this expression goes to (as $f \in C^1(U, B)$)

$$D_{(a_1,\ldots,a_k)} f(0, \ldots, a_l', \ldots, 0)$$

which is c l $/\mathbb{K}$ as function of a_l'. Hence, by Lemma (5.1.4), we have that $g^{(\widehat{a_l})}$ d a_l and

$$D_{a_l} g_l^{(\widehat{a_l})}(a_l') = D_{(a_1,\ldots,a_k)} f(0, \ldots, a_l', \ldots, 0).$$

Because $f \in C^1(U, B)$ we see that this expression is c as a function of $(a_1, \ldots, a_k, a_l')$.

Now Definition (5.1.12) gives

$$\frac{\partial f(a_1, \ldots, a_k)}{\partial a_l}(a_l') = D_{a_l} g_l^{(\widehat{a_l})}(a_l')$$

which yields the desired continuity of the partial derivatives of f.

We also obtain from $D_{(a_1,\ldots,a_k)} f$ l that,

$$\begin{aligned} D_{(a_1,\ldots,a_k)} f(a_1', \ldots, a_k') &= \sum_{l=1}^{k} D_{(a_1,\ldots,a_k)} f(0, \ldots, a_l', \ldots, 0) \\ &= \sum_{l=1}^{k} D_{a_l} g_l^{(\widehat{a_l})}(a_l') \\ &= \sum_{l=1}^{k} \frac{\partial f(a_1, \ldots, a_k)}{\partial a_l}(a_l') \end{aligned}$$

which gives Equation (5.2). □

Definition 5.1.14: Diffeomorphism

Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, $V \subseteq B$ open, and $k \in \mathbb{N}$.

Then a map $f : U \to V$ is called a C^k *diffeomorphism* if f is bijective, $f \in C^k(U, B)$, and $f^{-1} \in C^k(V, A)$.

If there exists a C^k diffeomorphism $f : U \to V$, then we call U and V C^k *diffeomorphic*.

Lemma 5.1.15
Let A, B $/\mathbb{K}$, $U \subseteq A$ open, and $V \subseteq B$ open.

Suppose $f : U \to V$ is a C^k diffeomorphism, then for all $a \in U$ we have

$$[D_a f]^{-1} = D_{f(a)} f^{-1}.$$

In particular for all $a \in U$, $D_a f : A \to B$ is a -isomorphism.

Proof. Let $f : U \to V$ be a C^k diffeomorphism and $a \in U$. Then $f^{-1} \circ f = \mathrm{id}_U$, so by Theorem (5.1.8) we have for all $a_1 \in A$ that $a_1 = D_a \,\mathrm{id}_U(a_1) = D_a(f^{-1} \circ f)(a_1) = D_{f(a)} f^{-1}(D_a f(a_1))$. Similarly, by considering $f \circ f^{-1} = \mathrm{id}_V$ for all $b_1 \in B$, $b_1 = D_a f(D_{f(a)} f^{-1}(b_1))$. Hence $[D_a f]^{-1} = D_{f(a)} f^{-1}$. □

A converse to Lemma (5.1.15) is given in Theorem (5.5.8).

Theorem 5.1.16: Symmetry of higher order derivatives
Let A, B $/\mathbb{K}$, $k \in \mathbb{N}$, $U \subseteq A$ open.

Then for any $f \in C^k(U, B)$ we have for all $a \in U$ and $a_1, \ldots, a_k \in A$ that for any $\pi \in S^k$,

$$D_a^k f(a_{\pi(1)}, \ldots, a_{\pi(k)}) = D_a^k f(a_1, \ldots, a_k).$$

Proof. Suppose $k = 2$. Let $f \in C^2(U, B)$, $a \in U$, $a_1, a_2 \in A$. Let $g \in B'$ be arbitrary.

The function $\mathbb{K}^2 \to A : (\alpha, \beta) \mapsto a + \alpha\, a_1 + \beta\, a_2$ and U is an open neighbourhood of a. Hence there exists a $\delta \in]0, \infty[$ such that for all $\alpha, \beta \in B_{\mathbb{K}}(0, \delta)$ we have $a + \alpha\, a_1 + \beta\, a_2 \in U$. Let $h : B_{\mathbb{K}}(0, \delta) \times B_{\mathbb{K}}(0, \delta) \to \mathbb{K}$ be given by $h(\alpha, \beta) := g(f(a + \alpha\, a_1 + \beta\, a_2))$. Then as g we have by Theorem (5.1.8) and Corollary (5.1.7) that

$$\begin{aligned}
\frac{\partial^2 h(\alpha, \beta)}{\partial \alpha\, \partial \beta} &= \frac{\partial}{\partial \alpha}\Big(D_{f(a+\alpha\, a_1+\beta\, a_2)} g\Big(D_{a+\alpha\, a_1+\beta\, a_2} f(a_2)\Big)\Big) \\
&= \frac{\partial}{\partial \alpha}\Big(g\Big(D_{a+\alpha\, a_1+\beta\, a_2} f(a_2)\Big)\Big) \\
&= \ldots \text{follow the same procedure} \ldots \\
&= g\Big(D^2_{a+\alpha\, a_1+\beta\, a_2} f(a_2, a_1)\Big)
\end{aligned}$$

and similar expressions for $\frac{\partial^2 h(\alpha,\beta)}{\partial \beta\, \partial \alpha}$, $\frac{\partial^2 h(\alpha,\beta)}{\partial \alpha\, \partial \alpha}$, and $\frac{\partial^2 h(\alpha,\beta)}{\partial \beta\, \partial \beta}$. Therefore, as $f \in C^2(U, B)$, h is twice partially differentiable map $\mathbb{K}^2 \to \mathbb{K}$ with continuous partial derivatives, therefore, from analysis on $\mathbb{K}$, we know that

$$g\Big(D^2_{a+\alpha\, a_1+\beta\, a_2} f(a_2, a_1)\Big) = \frac{\partial^2 h(\alpha, \beta)}{\partial \alpha\, \partial \beta} = \frac{\partial^2 h(\alpha, \beta)}{\partial \beta\, \partial \alpha} = g\Big(D^2_{a+\alpha\, a_1+\beta\, a_2} f(a_1, a_2)\Big).$$

As this is true for all $g \in A'$ we find at $\alpha = \beta = 0$ with Theorem (4.5.14) that

$$D_a^2 f(a_2, a_1) = D_a^2 f(a_1, a_2)$$

for all $a_1, a_2 \in A$.

For $k > 2$ the proof follows likewise (construct a function $h(\alpha_1, \ldots, \alpha_k) = g(f(a + \alpha_1 a_1 + \ldots + \alpha_k a_k))$ for $g \in B'$ and proceed in the same way). □

5.2 Multilinear families

In the treatment of [Chr1869] we will encounter a lot of functions that are multilinear in all but one of their variables. Therefore we will investigate such functions in this section.

Lemma 5.2.1: Differentiability of families of k- maps
Let $k \in \mathbb{N}$, A, B_1, …, B_k, C $/\mathbb{K}$, $U \subseteq A$ open.

Let $f : U \times B_1 \times \ldots \times B_k \to C$, such that for all $a \in U$ the map

$$B_1 \times \ldots \times B_k \to C : (b_1, \ldots, b_k) \mapsto f(a, b_1, \ldots, b_k)$$

is k- .

Then $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$ for some $l \in \mathbb{N}$ if and only if f and for all $b_1 \in B_1$, …, $b_k \in B_k$ the map

$$U \to C : a \mapsto f(a, b_1, \ldots, b_k)$$

is an element of $C^l(U, C)$.

In particular if this condition is satisfied we have for all $a \in U$, $a' \in A$, $b_1, b_1' \in B_1$, …, $b_k, b_k' \in B_k$ that

$$\begin{aligned} D_{(a,b_1,\ldots,b_k)} f(a', b_1', \ldots, b_k') = D_{(a,b_1,\ldots,b_k)} f(a', 0, \ldots, 0) \\ + \sum_{m=1}^{k} f(a, b_1, \ldots, b_{m-1}, b_m', b_{m+1}, \ldots, b_k). \end{aligned}$$

Proof. Let $f : U \times B_1 \times \ldots \times B_k \to C$ satisfy the conditions of the theorem.

If $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$ for $l \in \mathbb{N}$, then the condition is satisfied directly.

Suppose conversely that f and for all $b_1 \in B_1$, …, $b_k \in B_k$ we have $(U \to C : a \mapsto f(a, b_1, \ldots, b_k)) \in C^l(U, C)$. Fix $b_1 \in B_1$, …, $b_k \in B_k$ and denote

$$g : U \to C : a \mapsto f(a, b_1, \ldots, b_k),$$

then $g \in C^l(U, C)$ by assumption. Let $a' \in A$, $b_1' \in B_1$, …, $b_k' \in B_k$, then for $\alpha \in \mathbb{K}$ small enough but nonzero, using k-linearity

$$\begin{aligned} &\frac{1}{\alpha}\Big(f(a + \alpha\, a', b_1 + \alpha\, b_1', \ldots, b_k + \alpha\, b_k') - f(a, b_1, \ldots, b_k)\Big) \\ &= \frac{1}{\alpha}\Big(f(a + \alpha\, a', b_1, \ldots, b_k) \\ &\quad + \alpha \sum_{m=1}^{k} f(a + \alpha\, a', b_1 + \alpha\, b_1', \ldots, b_{m-1} + \alpha\, b_{m-1}', b_m', b_{m+1} + \alpha\, b_{m+1}', \ldots, b_k + \alpha\, b_k') \\ &\quad - f(a, b_1, b_2, \ldots, b_k)\Big) \\ &= \frac{1}{\alpha}\Big(g(a + \alpha\, a') - g(a)\Big) \\ &\quad + \sum_{m=1}^{k} f(a + \alpha\, a', b_1 + \alpha\, b_1', \ldots, b_{m-1} + \alpha\, b_{m-1}', b_m', b_{m+1} + \alpha\, b_{m+1}', \ldots, b_k + \alpha\, b_k'). \end{aligned}$$

Since f c and g d, we can take the limit $(\alpha, a', b_1', \ldots, b_k') \to (0, a'', b_1'', \ldots, b_k'')$ of this expression to obtain

$$D_a g(a'') + \sum_{m=1}^{k} f(a + 0, b_1 + 0, \ldots, b_{m-1} + 0, b_m'', b_{m+1} + 0, \ldots, b_k + 0).$$

Since this expression depends continuously on a, b_1, $\ldots$, b_k, a'', b_1'', $\ldots$, b_k'' because $g \in C^l(U, C)$ and f c, we see with Lemma (5.1.4) that $f \in C^1(U \times B_1 \times \ldots \times B_k, C)$ with derivative precisely given by this expression.

We can take further derivatives of our expression for Df and use induction to obtain that $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$. □

Lemma (5.2.1) shows that the derivative of a collection of k-l maps depending on a parameter $a \in U$, really only depends on the derivative with respect to a. This motivates a less cumbersome notation.

Definition 5.2.2: k-linear family
Let $k \in \mathbb{N}$, A, B_1, $\ldots$, B_k, C TVS $/\mathbb{K}$ T2 LC, $U \subseteq A$ open.

Then we call a map $f : U \times B_1 \times \ldots \times B_k \to C$ *a family of k-linear maps* or *a k-linear family* if for all $a \in U$ the map f_a defined by

$$f_a : B_1 \times \ldots \times B_k \to C : (b_1, \ldots, b_k) \mapsto f(a, b_1, \ldots, b_k)$$

is k-l.

Furthermore, if $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$ we denote for $a \in U$, $a_1, \ldots, a_l \in A$, $b_1 \in B_1$, $\ldots$, $b_k \in B_k$

$$D_a^l f(a_1, \ldots, a_l)(b_1, \ldots, b_k) := D_{(a, b_1, \ldots, b_k)}^l f((a_1, 0, \ldots, 0), \ldots, (a_l, 0, \ldots, 0)).$$

Note that by Lemma (5.2.1), the expression $D_a^l f(a_1, \ldots, a_l)(b_1, \ldots, b_k)$ completely determines the derivative of f, since

$$\begin{aligned} D_{(a, b_1, \ldots, b_k)} f(a', b_1', \ldots, b_k') = {} & D_a f(a')(b_1, \ldots, b_k) \\ & + \sum_{l=1}^{k} f_a(b_1, \ldots, b_{l-1}, b_l', b_{l+1}, \ldots, b_k). \end{aligned} \tag{5.3}$$

Example 5.2.3
Let $k \in \mathbb{N}$ and consider the *inner product*

$$\langle \cdot, \cdot \rangle : \mathbb{K}^k \times \mathbb{K}^k \to \mathbb{K}$$

defined by

$$\langle x, y \rangle := \sum_{l=1}^{k} x_l \, y_l$$

then $\langle \cdot, \cdot \rangle$ (considered as a map $\{0\} \times \mathbb{K}^k \times \mathbb{K}^k \to \mathbb{K}$) is a constant family of 2-linear maps.

Note that in particular for $k = 1$ the real and complex products fall in this category.

Theorem 5.2.4: Product rule

Let $k \in \mathbb{N}$, A, B_1, ..., B_k, C TVS $/\mathbb{K}$ T2 LC, $U \subseteq A$ open.

Let $f \in C^1(U \times B_1 \times \ldots \times B_k, C)$ be a family of k-linear maps.

Let $g_1 : U \to B_1$, ..., $g_k : U \to B_k$ and fix $a \in U$. If for all $1 \leq l \leq k$ we have that g_l d a, then the map

$$h : U \to C : a' \mapsto f_{a'}(g_1(a'), \ldots, g_k(a'))$$

is d a and

$$\begin{aligned} D_a h(a_1) &= D_a f(a_1)(g_1(a), \ldots, g_k(a)) \\ &\quad + \sum_{l=1}^{k} f_a\left(g_1(a), \ldots, g_{l-1}(a), D_a g_l(a_1), g_{l+1}(a), \ldots, g_k(a)\right). \end{aligned} \tag{5.4}$$

In particular if for a certain $l \in \mathbb{N}$, $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$, $g_1 \in C^l(U, B_1)$, ..., $g_k \in C^l(U, B_k)$, then $h \in C^l(U, C)$.

Proof. Consider the function $i : U \to U \times B_1 \times \ldots \times B_k : a' \mapsto (a', g_1(a'), \ldots, g_k(a'))$, then i d a and for $a'' \in A$ we have (as $a' \mapsto a'$ is c l, and g_1, ..., g_k d a, use Lemma (5.1.11))

$$D_a i(a'') = (a'', D_a g_1(a''), \ldots, D_a g_k(a'')).$$

Now $h(a) = f(i(a))$, f d $i(a)$, and i d a, so by Theorem (5.1.8) $h = f \circ i$ d a and

$$\begin{aligned} D_a h(a') &= D_{i(a)} f(D_a i(a')) \\ &= D_{(a, g_1(a), \ldots, g_k(a))} f(a', D_a g_1(a'), \ldots, D_a g_k(a')) \\ &\overset{(5.3)}{=} D_a f(a')(g_1(a), \ldots, g_k(a)) \\ &\quad + \sum_{l=1}^{k} f_a(g_1(a), \ldots, g_{l-1}(a), D_a g_l(a'), g_l(a), \ldots, g_k(a)). \end{aligned}$$

From Equation (5.4) we see that if $g_1 \in C^1(U, B_1)$, ..., $g_k \in C^1(U, B_k)$, then $D_a h(a')$ depends continuously on a and a' and hence $h \in C^1(U, C)$. In particular if for $l \in \mathbb{N}$, $f \in C^l(U \times B_1 \times \ldots \times B_k, C)$, $g_1 \in C^l(U, B_1)$, ..., $g \in C^l(U, B_k)$, then taking derivatives of Equation (5.4) and using induction, we see that $h \in C^l(U, C)$. □

Example 5.2.5

For the map $\langle \cdot, \cdot \rangle$ from Example (5.2.3) we see that for any two paths $f, g \in C^1(\mathbb{K}, \mathbb{K}^k)$, their inner product

$$h : \mathbb{K} \to \mathbb{K}, \qquad h(x) := \langle f(x), g(x) \rangle,$$

is C^1 and satisfies by Equation (5.4)

$$h'(x) = 0 + \langle f'(x), g(x) \rangle + \langle f(x), g'(x) \rangle.$$

Theorem 5.2.6: Differentiability of families of inverses
Let A, B, C TVS$/\mathbb{K}$ T2 LC, $U \subseteq A$ open, $k \in \mathbb{N}$.

Let $f \in C^k(U \times B, C)$ be a family of linear maps such that for all $a \in U$, $f_a : B \to C$ is bijective. Denote the family of inverses by $g : U \times C \to B : (a, c) \mapsto f_a^{-1}(c)$.

If g c, then $g \in C^k(U \times C, B)$.

In particular, for all $a \in U$, $a_1 \in A$, $c \in C$ we have

$$D_a g(a_1)(c) = -g_a(D_a f(a_1)(g_a(c))). \tag{5.5}$$

Proof. Let $a \in U$, $a_1 \in A$, $c, c_1 \in C$, and $\alpha \in \mathbb{K}$ small enough (A LC), then using linearity and invertibility

$$\begin{aligned}
&\frac{1}{\alpha}\Big(g(a + \alpha\, a_1, c + \alpha\, c_1) - g(a, c)\Big) \\
&= \frac{1}{\alpha}\Big(g(a + \alpha\, a_1, f(a, g(a, c + \alpha\, c_1))) - g(a + \alpha\, a_1, f(a + \alpha\, a_1, g(a, c)))\Big) \\
&= -g\Big(a + \alpha\, a_1, \frac{1}{\alpha}\Big(f(a + \alpha\, a_1, g(a, c)) - f(a, g(a, c + \alpha\, c_1))\Big)\Big) \\
&= -g\Big(a + \alpha\, a_1, \frac{1}{\alpha}\Big(f(a + \alpha\, a_1, g(a, c)) - f(a, g(a, c))\Big)\Big) \\
&\quad + g(a + \alpha\, a_1, \frac{1}{\alpha}\, \alpha\, f(a, g(a, c_1))) \\
&= -g\Big(a + \alpha\, a_1, \frac{1}{\alpha}\Big(f(a + \alpha\, a_1, g(a, c)) - f(a, g(a, c))\Big)\Big) \\
&\quad + g(a + \alpha\, a_1, c_1).
\end{aligned}$$

Since all involved functions depend continuously on α, a_1, and c_1, and f d $(a, g(a, c))$ we therefore find with Lemma (5.1.4) that g d (a, c) and

$$D_{(a,c)} g(a_1, c_1) = -g_a(D_{(a, g_a(c))} f(a_1, 0)) + g(a, c_1)$$

which becomes Equation (5.5) in the notation of Definition (5.2.2). Because g is continuous and f continuously differentiable, $D_{(a,c)} g(a_1, c_1)$ depends continuously on a, c, a_1, and c_1, and therefore $g \in C^1(U \times C, B)$.

Now we can use induction, the composition rule from Theorem (5.1.8) and product rule from Theorem (5.2.4) to find that $g \in C^k(U \times C, B)$ because $f \in C^k(U \times B, C)$, by taking derivatives of Equation (5.5). □

5.3 Integration

Definition 5.3.1: Partitions of an interval
Let $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$.

Then a *partition of the interval* $[\alpha, \beta]$ is a collection $\gamma_0, \ldots, \gamma_k \in \mathbb{R}$ satisfying $\alpha = \gamma_0 < \gamma_1 < \ldots < \gamma_k = \beta$.

For any two partitions $\gamma_0, \ldots \gamma_k$ and $\delta_0, \ldots, \delta_l$ of $[\alpha, \beta]$ we say that $\delta_0, \ldots, \delta_l$ *is a refinement of* $\gamma_0, \ldots, \gamma_k$ if there exist integers $i_0, \ldots, i_k \in \mathbb{N}$ such that $\gamma_j = \delta_{i_j}$ for all $1 \le j \le k$.

We will now introduce the notion of an integral, not in terms of measuring sets (i.e. measuring their volume, as is done with Lebesgue integration), but as an inverse operation to differentiation as presented in Section 5.1. In this section we demand that A UC to ensure that the integrals actually exist.

Definition 5.3.2: Integral
Let A Vs /$\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval and $f : S \to A$ a map.

Let $\alpha, \beta \in S$, $\alpha < \beta$. Then we say for $a \in A$ that *the integral of f over $[\alpha, \beta]$ equals a*, denoted by

$$\int_\alpha^\beta f = a,$$

if for all $i \in I$ and $\epsilon \in]0, \infty[$ there exist a partition $\gamma_0, \ldots, \gamma_k$ of $[\alpha, \beta]$, such that for all refinements $\delta_0, \ldots, \delta_l$ of this partition we have that

$$\left\| \Big[\sum_{m=0}^{l-1} (\delta_{m+1} - \delta_m)\, f\Big(\frac{\delta_{m+1} + \delta_m}{2}\Big) \Big] - a \right\|_i < \epsilon.$$

If it is not the case that $\int_\alpha^\beta f = a$, we write $\int_\alpha^\beta f \neq a$.

For $\alpha > \beta$ we define

$$\int_\alpha^\beta f := -\int_\beta^\alpha f$$

whenever the latter exists and for $\alpha = \beta$ we define

$$\int_\alpha^\alpha f := 0.$$

Example 5.3.3
Let A Vs /$\mathbb{K}$ T2 LC UC, $a \in A$, $S \subseteq \mathbb{R}$ an open interval and $f : S \to A$ the constant map $f(\alpha) = a$ for all $\alpha \in S$.

Then for any $\alpha, \beta \in S$, $\alpha < \beta$ and any partition $\gamma_0, \ldots, \gamma_k$ of $[\alpha, \beta]$ we have $\sum_{l=0}^{k-1} (\gamma_{l+1} - \gamma_l)\, f((\gamma_{l+1} + \gamma_l)/2) = \sum_{l=0}^{k-1} (\gamma_{l+1} - \gamma_l)\, a = (\beta - \alpha)\, a$. Therefore, for any $\alpha, \beta \in S$ (swap them if $\alpha > \beta$),

$$\int_\alpha^\beta (\gamma \mapsto a) = (\beta - \alpha)\, a.$$

Lemma 5.3.4
Let A Vs /$\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval, and $f : S \to A$ a map.

Suppose f C, then for any $\alpha, \beta \in S$ there exists a unique $a \in A$ such that $\int_\alpha^\beta f = a$.

Proof. Let $\alpha, \beta \in S$, $\alpha < \beta$. Create for each $k \in \mathbb{N}$ the partition $\gamma_0^{(k)}, \ldots, \gamma_{2^k}^{(k)}$ of $[\alpha, \beta]$ by dividing $[\alpha, \beta]$ into 2^k equal pieces: define $\gamma_l^{(k)} := \alpha + \frac{\beta - \alpha}{2^k}\, l$ for $0 \le l \le 2^k$. Note that $\gamma_l^{(k)} = \gamma_{2l}^{(k+1)}$ by definition.

Now construct the sequence $x : \mathbb{N} \to A$ by defining for all $k \in \mathbb{N}$

$$x_k := \sum_{l=0}^{2^k} (\gamma_{l+1}^{(k)} - \gamma_l^{(k)})\, f\Big(\frac{\gamma_{l+1}^{(k)} + \gamma_l^{(k)}}{2}\Big).$$

Let $i \in I$, $\epsilon \in]0, \infty[$ be given, then ($[\alpha, \beta]$ Cpt, f C, use Theorem (2.5.23)) there exists a $\delta \in]0, \infty[$ such that $\|f(\alpha_1) - f(\alpha_2)\|_i < \frac{\epsilon}{\beta - \alpha}$ whenever $|\alpha_1 - \alpha_2| < \delta$. Pick $k_0 \ge 1 + \lceil {}^2\log\left(\frac{\beta - \alpha}{\delta}\right) \rceil$, then for any $k \ge k_0$ we have that $\gamma_{l+1}^{(k)} - \gamma_l^{(k)} =$

$\frac{\beta-\alpha}{2^k} < \delta$. Let $k \geq k_0$ be arbitrary and $\delta_0, \dots, \delta_l$ any refinement of $\gamma_0^{(k)}, \dots, \gamma_{2^k}^{(k)}$. Then there exist $i_0, \dots, i_{2^k} \in \mathbb{N}$ such that $\gamma_m^{(k)} = \delta_{i_m}$ for $0 \leq m \leq 2^k$. So,

$$\begin{aligned}
&x_k - \sum_{m=0}^{l} (\delta_{m+1} - \delta_m)\, f((\delta_{m+1} + \delta_m)/2) \\
&= \sum_{m=0}^{2^k} (\gamma_{m+1}^{(k)} - \gamma_m^{(k)})\, f((\gamma_{m+1}^{(k)} + \gamma_m^{(k)})/2) \\
&- \sum_{n=0}^{2^k} \sum_{o=i_n}^{i_{n+1}-1} (\delta_{o+1} - \delta_o)\, f((\delta_{o+1} + \delta_o)/2) \\
&= \sum_{m=0}^{2^k} \left(\sum_{p=i_m}^{i_{m+1}-1} (\delta_{p+1} - \delta_p) \right) f((\gamma_{m+1}^{(k)} + \gamma_m^{(k)})/2) \\
&- \sum_{n=0}^{2^k} \sum_{o=i_n}^{i_{n+1}-1} (\delta_{o+1} - \delta_o)\, f((\delta_{o+1} + \delta_o)/2) \\
&= \sum_{m=0}^{2^k} \sum_{n=i_m}^{i_{m+1}-1} (\delta_{n+1} - \delta_n) \Big(f((\gamma_{m+1}^{(k)} + \gamma_m^{(k)})/2) - f((\delta_{n+1} + \delta_n)/2) \Big).
\end{aligned}$$

With this,

$$\begin{aligned}
&\left\| \sum_{m=0}^{l} (\delta_{m+1} - \delta_m)\, f((\delta_{m+1} + \delta_m)/2) - x_k \right\|_i \\
&= \left\| \sum_{m=0}^{2^k} \sum_{n=i_m}^{i_{m+1}-1} (\delta_{n+1} - \delta_n) \Big(f((\gamma_{m+1}^{(k)} + \gamma_m^{(k)})/2) - f((\delta_{n+1} + \delta_n)/2) \Big) \right\|_i \\
&\leq \sum_{m=0}^{2^k} \sum_{n=i_m}^{i_{m+1}-1} |\delta_{n+1} - \delta_n| \, \Big\| f((\gamma_{m+1}^{(k)} + \gamma_m^{(k)})/2) - f((\delta_{n+1} + \delta_n)/2) \Big\|_i \\
&< \sum_{m=0}^{2^k} \sum_{n=i_m}^{i_{m+1}-1} (\delta_{n+1} - \delta_n) \frac{\epsilon}{\beta - \alpha} = \epsilon
\end{aligned}$$

where we used uniform continuity of f together with the fact that for each $0 \leq m \leq 2^k$ and $i_m \leq n < i_{m+1}$ we have $[\delta_n, \delta_{n+1}] \subseteq [\gamma_m^{(k)}, \gamma_{m+1}^{(k)}]$ and $0 < \gamma_{m+1}^{(k)} - \gamma_m^{(k)} < \delta$.

In particular, for all $k, l \geq k_0$ we have $\|x_k - x_l\|_i < \epsilon$. Since this is true for all $i \in \mathbb{N}$ (A UC), there exists a unique (A T2) limit $a \in A$ of the sequence x.

Also, let $i \in I$, $\epsilon \in]0, \infty[$ be given and pick k_0 at least as large as before, but also such that $\|x_k - a\|_i < \epsilon$ for all $k \geq k_0$ (use $\lim_{k\to\infty} x_k = a$). Then for any refinement $\delta_0, \ldots, \delta_l$ of $\gamma_0^{(k_0)}, \ldots, \gamma_{2^{k_0}}^{(k_0)}$ we have by the above

$$\begin{aligned}
&\left\| \sum_{m=0}^{l} (\delta_{m+1} - \delta_m)\, f((\delta_{m+1} + \delta_m)/2) - a \right\|_i \\
&\leq \left\| \sum_{m=0}^{l} (\delta_{m+1} - \delta_m)\, f((\delta_{m+1} + \delta_m)/2) - x_{k_0} \right\|_i + \|x_{k_0} - a\|_i \\
&< \epsilon + \epsilon = 2\,\epsilon.
\end{aligned}$$

Therefore $\int_\alpha^\beta f = a$.

If $\alpha = \beta$ then we have by definition $\int_\alpha^\beta f = \int_\alpha^\alpha f = 0 \in A$.

If $\alpha > \beta$, then by the above there exists an $a \in A$ such that $\int_\beta^\alpha f = a$. So $\int_\alpha^\beta f = -\int_\beta^\alpha f = -a \in A$ exists by definition. ☐

Theorem 5.3.5: Integration
Let A Vs $/\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval, and $f : S \to A$ C.

- For any $\gamma \in \mathbb{K}$, $g : S \to A$ C, and $\alpha, \beta \in S$ we have
$$\int_\alpha^\beta (f + \gamma\, g) = \int_\alpha^\beta f + \gamma \int_\alpha^\beta g.$$
- For any $\alpha, \beta, \gamma \in S$ we have
$$\int_\alpha^\gamma f = \int_\alpha^\beta f + \int_\beta^\gamma f.$$
- For any seminorm $\|\cdot\| : A \to \mathbb{R}$ C and $\alpha, \beta \in S$ we have
$$\left\| \int_\alpha^\beta f \right\| \leq \left| \int_\alpha^\beta \|f\| \right|.$$
- For all $\alpha, \beta \in S$ we have for any B Vs $/\mathbb{K}$ T2 LC UC, $g \in A \to B$ C l that
$$g\Big(\int_\alpha^\beta f \Big) = \int_\alpha^\beta (g \circ f).$$

Proof. In this proof we repeatedly use the fact that A T2 which implies that limits (in particular the limit of the sequence x from the proof of Lemma (5.3.4)) are unique.

- Let $\gamma \in \mathbb{K}$, $g : S \to A$ C. Suppose $\alpha < \beta$. Then $f + \gamma\, g$ C, so there exists a unique value for $\int_\alpha^\beta (f + \gamma\, g)$ by Lemma (5.3.4). Since for any partition $\gamma_0, \ldots, \gamma_k$ of $[\alpha, \beta]$ we have $\sum_{l=0}^{k} (\gamma_{l+1} - \gamma_l)\, (f + \gamma\, g)((\gamma_{l+1} + \gamma_l)/2) = \sum_{l=0}^{k} (\gamma_{l+1} - \gamma_l)\, f((\gamma_{l+1} + \gamma_l)/2) + \gamma \sum_{m=0}^{k} (\gamma_{m+1} - \gamma_m)\, g((\gamma_{m+1} + \gamma_m)/2)$, this unique value necessarily equals $\int_\alpha^\beta f + \gamma \int_\alpha^\beta g$. For the case that $\alpha = \beta$ we obtain $0 = 0 + \gamma\, 0$ and when $\alpha > \beta$ we can swap α and β and add minus signs to the left and right hand sides of $\int_\alpha^\beta (f + \gamma g) = \int_\alpha^\beta f + \gamma \int_\alpha^\beta g$ to obtain the desired result.

- Let $\alpha, \beta, \gamma \in S$, $\alpha < \beta < \gamma$. Then by Lemma (5.3.4), $\int_\alpha^\gamma f$, $\int_\alpha^\beta f$, $\int_\beta^\gamma f$ exist in A. Furthermore, let $\gamma_0, \ldots, \gamma_k$ be any partition of $[\alpha, \gamma]$, then by possibly refining, through adding β to this partition, we may suppose that $\gamma_l = \beta$ (and conversely, we can concatenate any two partitions of $[\alpha, \beta]$ and $[\beta, \gamma]$ to obtain a partition of $[\alpha, \gamma]$). Then $\sum_{m=0}^{k}(\gamma_{m+1} - \gamma_m)\, f((\gamma_{m+1} + \gamma_m)/2) = \sum_{m=0}^{l-1}(\gamma_{m+1} - \gamma_m)\, f((\gamma_{m+1} + \gamma_m)/2) + \sum_{n=l}^{k}(\gamma_{n+1} - \gamma_n)\, f((\gamma_{n+1} + \gamma_n)/2)$. As both terms on the right hand side will tend to $\int_\alpha^\beta f$ and $\int_\beta^\gamma f$ for increasingly finer partitions, necessarily $\int_\alpha^\gamma f = \int_\alpha^\beta f + \int_\beta^\gamma f$.

  For other configurations, instead of $\alpha < \beta < \gamma$, simply apply minus signs at the appropriate positions and reduce to the $\cdot \le \cdot \le \cdot$ case.

- Let $\|\cdot\| : A \to \mathbb{R}$ c be a seminorm. Then f c and $\|f\| := \|\cdot\| \circ f$ c, so by Lemma (5.3.4), both $\int_\alpha^\beta f$ and $\int_\alpha^\beta \|f\|$ exist in A and $\mathbb{R}$ respectively and are unique. Suppose $\alpha < \beta$, the result now follows from repeated applications of the triangle inequality and continuity of $\|\cdot\|$: for any partition $\gamma_0, \ldots, \gamma_k$ of $[\alpha, \beta]$ we have $\|\sum_{l=0}^{k}(\gamma_{l+1} - \gamma_l)\, f((\gamma_{l+1} + \gamma_l)/2)\| \le \sum_{l=0}^{k} |\gamma_{l+1} - \gamma_l|\, \|f((\gamma_{l+1} + \gamma_l)/2)\| = \sum_{l=0}^{k}(\gamma_{l+1} - \gamma_l)\, \|f\|((\gamma_{l+1} + \gamma_l)/2)$. Hence necessarily $\|\int_\alpha^\beta f\| \le \int_\alpha^\beta \|f\| = |\int_\alpha^\beta \|f\||$ since $\alpha < \beta$ and therefore $\int_\alpha^\beta \|f\| \ge 0$ because all terms in the sums converging to this integral are positive. The case $\alpha = \beta$ is direct ($0 \le 0$) and for $\alpha > \beta$ we have by the above $\|\int_\beta^\alpha f\| \le \int_\beta^\alpha \|f\|$, so $\|\int_\alpha^\beta f\| = \|-\int_\beta^\alpha f\| = \|\int_\beta^\alpha f\| \le \int_\beta^\alpha \|f\| = -\int_\alpha^\beta \|f\| = |\int_\alpha^\beta \|f\||$ as $\int_\beta^\alpha \|f\| \ge 0$.

- Suppose $\alpha < \beta$. Let $g \in A \to B$ c l, then $g \circ f : S \to B$ c, so both $\int_\alpha^\beta f$ and $\int_\alpha^\beta (g \circ f)$ exist and are unique. As g l, we have for any partition $\gamma_0, \ldots, \gamma_k$ of $[\alpha, \beta]$ that $g\Big(\sum_{l=0}^{k}(\gamma_{l+1} - \gamma_l) f((\gamma_{l+1} + \gamma_l)/2)\Big) = \sum_{l=0}^{k}(\gamma_{l+1} - \gamma_l)(g \circ f)((\gamma_{l+1} + \gamma_l)/2)$, so using continuity of g we find that necessarily $g(\int_\alpha^\beta f) = \int_\alpha^\beta (g \circ f)$.

  The cases $\alpha = \beta$ and $\alpha > \beta$ follow directly.

□

Example 5.3.6: Integration on $\mathbb{R}$ and $\mathbb{C}$

Let A Vs $/\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval, $f : S \to A$ c, and $\alpha, \beta \in S$.

Then for any $g \in A'$ we have by Theorem (5.3.5) (as $g : A \to \mathbb{K}$ c l $/\mathbb{K}$) that $g(\int_\alpha^\beta f) = \int_\alpha^\beta (g \circ f)$. However $g \circ f : S \to \mathbb{K}$, so looking at Definition (5.3.2) we see that $\int_\alpha^\beta (g \circ f)$ and the Riemann integral $\int_\alpha^\beta (g \circ f)(x)\,\mathrm{d}x$ must agree, because both exist ($[\alpha, \beta]$ compact, $g \circ f$ c) and they satisfy the same limiting procedure. Therefore, for any $g \in A'$,

$$g\Big(\int_\alpha^\beta f\Big) = \int_\alpha^\beta (g \circ f)(x)\,\mathrm{d}x.$$

Lemma 5.3.7

Let A, B Vs $/\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval, $U \subseteq A$ open, and $f : S \times U \to B$ c.

Then for any $\alpha, \beta \in S$ the map $g : U \to B$ defined by

$$g(a) := \int_\alpha^\beta \Big(\gamma \mapsto f(\gamma, a)\Big)$$

is $\textcircled{c}$.

Furthermore, if for $k \in \mathbb{N}$, $f \in C^k(S \times U, B)$, then $g \in C^k(U, B)$ and

$$D_a g(a_1) = \int_\alpha^\beta \Big(\gamma \mapsto D_{(\gamma,a)} f(0, a_1)\Big).$$

Proof. Fix $a \in U$ and let $i \in I$, $\epsilon \in]0, \infty[$ be given. Let $\gamma \in [\alpha, \beta]$, then because $f \ \textcircled{c} (\gamma, a)$ there exists a $\delta_\gamma \in]0, \infty[$ and an open neighbourhood U_γ of a in U such that for all $\gamma_1 \in [\alpha, \beta]$, $a_1 \in U$ we have that $|\gamma_1 - \gamma| < \delta_\gamma$ and $a_1 \in U_\gamma$ together imply that $\|f(\gamma_1, a_1) - f(\gamma, a)\|_i < \epsilon/2$. The collection $\{]\gamma - \delta_\gamma, \gamma + \delta_\gamma[\subseteq S | \gamma \in [\alpha, \beta]\}$ forms an open cover of $[\alpha, \beta]$. Since $[\alpha, \beta]$ is compact there exists a finite number of $\gamma_1, \ldots, \gamma_k \in [\alpha, \beta]$ such that $[\alpha, \beta] \subseteq \bigcup_{l=1}^k]\gamma_l - \delta_{\gamma_l}, \gamma_l + \delta_{\gamma_l}[$. Choose $U_1 := \bigcap_{l=1}^k U_{\gamma_l}$ which is (finite intersection of open sets) an open neighbourhood of a in U. Let $\gamma \in [\alpha, \beta]$ and $a_1 \in U_1$ be arbitrary. Then there exists an $l \in \{1, \ldots, k\}$ such that $\gamma \in]\gamma_l - \delta_{\gamma_l}, \gamma_l + \delta_{\gamma_l}[$. Hence $|\gamma - \gamma_l| < \delta_{\gamma_l}$ and $a_1 \in U_1 \subseteq U_{\gamma_l}$, so $\|f(\gamma, a_1) - f(\gamma_l, a)\|_i < \epsilon/2$. Therefore

$$\begin{aligned} \|f(\gamma, a_1) - f(\gamma, a)\|_i &\le \|f(\gamma, a_1) - f(\gamma_l, a) + f(\gamma_l, a) - f(\gamma, a)\|_i \\ &\le \|f(\gamma, a_1) - f(\gamma_l, a)\|_i + \|f(\gamma_l, a) - f(\gamma, a)\|_i \\ &< \epsilon/2 + \epsilon/2 = \epsilon, \end{aligned}$$

as $a \in U_{\gamma_l}$. So for any $a \in U$, $i \in I$, $\epsilon \in]0, \infty[$ there exists an open neighbourhood U_1 of a in U such that for all $a_1 \in U_1$ and $\gamma \in [\alpha, \beta]$ we have $\|f(\gamma, a_1) - f(\gamma, a)\|_i < \epsilon$.

Hence, for any $a_1 \in U_1$ we have (use Theorem (5.3.5))

$$\begin{aligned} \|g(a_1) - g(a)\|_i &= \left\| \int_\alpha^\beta (\gamma \mapsto f(\gamma, a_1)) - \int_\alpha^\beta (\gamma \mapsto f(\gamma, a)) \right\|_i \\ &= \left\| \int_\alpha^\beta (\gamma \mapsto f(\gamma, a_1) - f(\gamma, a)) \right\|_i \\ &\le \left| \int_\alpha^\beta (\gamma \mapsto \|f(\gamma, a_1) - f(\gamma, a)\|_i) \right| \\ &\le \left| \int_\alpha^\beta \epsilon \right| = \epsilon\, (\beta - \alpha). \end{aligned}$$

Since this is true for all $a \in U$, $i \in I$, $\epsilon \in]0, \infty[$, we find that $g \ \textcircled{c}$.

Now suppose that $f \in C^k(S \times U, B)$. First of all note that for $\delta \in \mathbb{K}$ nonzero but small enough, $a \in U$, and $a_2 \in A$ we have by Theorem (5.3.5)

$$\begin{aligned} \frac{1}{\delta}\Big(g(a + \delta\, a_2) - g(a)\Big) &= \frac{1}{\delta}\Big(\int_\alpha^\beta \Big(\gamma \mapsto f(\gamma, a + \delta\, a_2)\Big) - \int_\alpha^\beta \Big(\gamma \mapsto f(\gamma, a)\Big)\Big) \\ &= \frac{1}{\delta}\Big(\int_\alpha^\beta \Big(\gamma \mapsto f(\gamma, a + \delta\, a_2) - f(\gamma, a)\Big)\Big) \\ &= \int_\alpha^\beta \Big(\gamma \mapsto \frac{1}{\delta}\Big(f(\gamma, a + \delta\, a_2) - f(\gamma, a)\Big)\Big). \end{aligned}$$

As $f \in C^1(S \times U, B)$, the function defined for all $\gamma \in S$, $\delta \neq 0$ small, and $a_2 \in A$ by $(\gamma, \delta, a_2) \mapsto \frac{1}{\delta}\Big(f(\gamma, a + \delta\, a_2) - f(\gamma, a)\Big)$ and for $\delta = 0$, $a_1 \in A$ by $(\gamma, 0, a_1) \mapsto D_{(\gamma,a)} f(0, a_1)$ is c. Therefore, by the first part of this lemma, so is $(\delta, a_2) \mapsto \frac{1}{\delta}\Big(g(a + \delta\, a_2) - g(a)\Big)$ and for $(\delta, a_2) \to (0, a_1)$, this function goes to $\int_\alpha^\beta \Big(\gamma \mapsto D_{(\gamma,a)} f(0, a_1)\Big)$. Hence g d a and $D_a g(a_1)$ is exactly given by this integral. Again using the first part of the lemma we see from the expression of $D_a g(a_1)$ that it depends continuously on a and a_1, therefore $g \in C^1(U, B)$. Use induction to obtain that $g \in C^k(U, B)$. □

Theorem 5.3.8: Fundamental theorem of integration
Let A TVS $/\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval, and $f : S \to A$ c.

Then for any $\alpha \in S$, the map $g_\alpha : S \to A$ defined by

$$g_\alpha(\beta) := \int_\alpha^\beta f$$

for all $\beta \in S$, satisfies $g_\alpha \in C^1(S, A)$ and $D_\beta g_\alpha(1) = f(\beta)$ for all $\beta \in S$.

On the other hand, for any $g \in C^1(S, A)$ satisfying $D_\alpha g(1) = f(\alpha)$ for all $\alpha \in S$, we have that for all $\alpha, \beta \in S$,

$$g(\beta) - g(\alpha) = \int_\alpha^\beta f.$$

Proof. In this proof we repeatedly use Theorem (5.3.5).

Fix $\alpha \in S$ and define $g_\alpha : S \to A$ as above. Let $\beta \in S$, $\epsilon \in]0, \infty[$, and $i \in I$ be given.

As f c β there exists a $\delta \in]0, \infty[$ such that for all $\gamma \in S - \beta$ satisfying $|\gamma| < \delta$ we have $\|f(\beta + \gamma) - f(\beta)\|_i < \epsilon/2$. Then for any $\gamma \in (S - \beta) \cap] - \delta, \delta[$,

$$g_\alpha(\beta + \gamma) - g_\alpha(\beta) = \int_\alpha^{\beta+\gamma} f - \int_\alpha^\beta f = \int_\beta^{\beta+\gamma} f.$$

So in the case that $\gamma \neq 0$, we have (use Example (5.3.3))

$$\begin{aligned}
\left\| \frac{1}{\gamma}\Big(g_\alpha(\beta + \gamma) - g_\alpha(\beta)\Big) - f(\beta) \right\|_i &= \left\| \frac{1}{\gamma}\Big(\int_\beta^{\beta+\gamma} f\Big) - \frac{(\beta + \gamma) - \beta}{\gamma} f(\beta) \right\|_i \\
&\leq \frac{1}{|\gamma|} \left| \int_\beta^{\beta+\gamma} \|f - f(\beta)\|_i \right| \\
&\leq \frac{1}{|\gamma|} \left| \int_\beta^{\beta+\gamma} \frac{\epsilon}{2} \right| \\
&= \frac{1}{|\gamma|} |\gamma| \frac{\epsilon}{2} < \epsilon.
\end{aligned}$$

Therefore (this is true for all $\epsilon \in]0, \infty[$, $i \in I$)

$$\lim_{\gamma \to 0} \frac{1}{\gamma}\Big(g_\alpha(\beta + \gamma 1) - g_\alpha(\beta)\Big) = f(\beta).$$

Hence by Corollary (5.1.7) g_α d β with $D_\beta g_\alpha(\gamma) = \gamma f(\beta)$ for all $\beta \in S$, $\gamma \in \mathbb{R}$. So g_α d S. The map $S \times \mathbb{R} \to A : (\beta, \gamma) \mapsto D_\beta g_\alpha(\gamma) = \gamma f(\beta)$ c since f c and scalar multiplication c. Hence $g_\alpha \in C^1(S, A)$.

Now let $g \in C^1(S, A)$ and suppose that for all $\alpha \in S$, $D_\alpha g(1) = f(\alpha)$. Fix some $\gamma \in S$, then in particular for all $\alpha \in S$ we have $D_\alpha g(1) = f(\alpha) = D_\alpha g_\gamma(1)$. Hence the map $g - g_\gamma$ (by Theorem (5.1.8)) satisfies $(g - g_\gamma) \in C^1(S, A)$ and $D_\alpha(g - g_\gamma) = D_\alpha g - D_\alpha g_\gamma = 0$. By Theorem (5.1.8) $g - g_\gamma$ is constant: there exists an $a \in A$ such that $g(\alpha) - g_\gamma(\alpha) = a$ for all $\alpha \in S$. Therefore $g(\beta) - g(\alpha) = g_\gamma(\beta) + a - g_\gamma(\alpha) - a = \int_\gamma^\beta f - \int_\gamma^\alpha f = \int_\alpha^\beta f$. $\square$

This immediately gives the following result, which is slightly counterintuitive in light of Example (5.1.10).

Corollary 5.3.9: Integration as antiderivative

Let A Vs $/\mathbb{K}$ T2 LC UC, $S \subseteq \mathbb{R}$ an open interval.

Let

$$C := \{f : S \to A \mid \exists a \in A : \forall \alpha \in S : f(\alpha) = a\}$$

be the collection of constant functions, considered as Vs $/\mathbb{K}$. Let $k \in \mathbb{N}$, $\alpha \in S$, then $C \leq C^{k+1}(S, A)$ and the maps

$$C^k(S, A) \to C^{k+1}(S, A)/C : f \mapsto \left[\beta \mapsto \int_\alpha^\beta f\right]$$
$$C^{k+1}(S, A)/C \to C^k(S, A) : [g] \mapsto \Big(\beta \mapsto D_\beta g(1)\Big)$$

are l $/\mathbb{K}$ and inverses of each other.

In particular as Vs $/\mathbb{K}$ we have for all $k \in \mathbb{N}$,

$$C^k(S, A) \simeq C^{k+1}(S, A)/A,$$

while at the same time it is also true that

$$C^k(S, A) \geq C^{k+1}(S, A).$$

Proof. Let $k \in \mathbb{N}$, $\alpha \in S$. By Theorem (2.1.28) and Theorem (5.1.8) we know that $C \leq C^{k+1}(S, A)$ as Vs $/\mathbb{K}$. Furthermore $C \simeq A$ as Vs $/\mathbb{K}$ via $A \to C : a \mapsto (\beta \mapsto a)$ and $C \to A : f \mapsto f(\alpha)$. The map $C^{k+1}(S, A)/C \to C^k(S, A)$ is well-defined, because for any two $g, h \in C^{k+1}(S, A)$ that differ by a constant $a \in A$ (so $g(\beta) = h(\beta) + a$ for all $\beta \in S$) we have (Theorem (5.1.8)) that $D_\beta g(1) = D_\beta h(1) + 0 = D_\beta h(1)$, so for all $h \in [g]$, $D_\beta h(1) = D_\beta g(1)$.

Let $f \in C^k(S, A)$ and choose $g : S \to A : \beta \mapsto \int_\alpha^\beta f$. Then by Theorem (5.3.8) the function $h : S \to A : \beta \mapsto D_\beta g(1)$ satisfies $h(\beta) = D_\beta g(1) = f(\beta)$ for all $\beta \in S$, so $h = f$.

Let $[g] \in C^{k+1}(S, A)/C$ and choose $f : S \to A : \beta \mapsto D_\beta g(1)$. Then by Theorem (5.3.8) the function $h : S \to A : \beta \mapsto \int_\alpha^\beta f$ satisfies $h(\beta) = \int_\alpha^\beta f = g(\beta) - g(\alpha)$, so $[h] = [g]$ since the function $\beta \mapsto -g(\alpha)$ is constant and hence an element of C.

So these two maps are inverses of each other and l $/\mathbb{K}$ by Theorem (5.1.8) and Theorem (5.3.5), hence

$$C^k(S, A) \simeq C^{k+1}(S, A)/C \simeq C^{k+1}(S, A)/A$$

as Vs $/\mathbb{K}$. From Example (5.1.10) we know that $C^{k+1}(S, A) \leq C^k(S, A)$. $\square$

Lemma 5.3.10: Partial integration
Let A /$\mathbb{K}$, $S \subseteq \mathbb{R}$ an open interval, $f \in C^1(S, \mathbb{K})$, and $g \in C^1(S, A)$.
Then for any $\alpha, \beta \in S$ we have

$$\int_\alpha^\beta (\gamma \mapsto D_\gamma f(1)\, g(\gamma)) = f(\beta)\, g(\beta) - f(\alpha)\, g(\alpha) - \int_\alpha^\beta (\gamma \mapsto f(\gamma)\, D_\gamma g(1)).$$

Proof. The scalar multiplication map $\mathbb{K} \times A \to A$ is 2- /$\mathbb{K}$ and f and g are C^1, so by Theorem (5.2.4) we obtain that the function

$$h : S \to A : \gamma \mapsto f(\gamma)\, g(\gamma)$$

satisfies $h \in C^1(S, A)$, and

$$D_\gamma h(1) = D_\gamma f(1)\, g(\gamma) + f(\gamma)\, D_\gamma g(1).$$

By Theorem (5.3.5) and Theorem (5.3.8) we therefore obtain that

$$\begin{aligned} f(\beta)g(\beta) - f(\alpha)g(\alpha) &= h(\beta) - h(\alpha) \\ &= \int_\alpha^\beta (\gamma \mapsto D_\gamma h(1)) \\ &= \int_\alpha^\beta (\gamma \mapsto D_\gamma f(1)g(\gamma) + f(\gamma)D_\gamma g(1)) \\ &= \int_\alpha^\beta (\gamma \mapsto D_\gamma f(1)g(\gamma)) + \int_\alpha^\beta (\gamma \mapsto f(\gamma)D_\gamma g(1)) \end{aligned}$$

which shows the desired result. □

We can now generalise the derivative as a local, linear approximation of a given function near a given point, to the Taylor sequence, which gives us a linear, quadratic, cubic, ... approximation.

Theorem 5.3.11: Taylor
Let A, B /$\mathbb{K}$, $U \subseteq A$ open abc, $k \in \mathbb{N}$, and $f \in C^{k+1}(U, B)$.
Then for any $a_1 \in U - a$ we have

$$\begin{aligned} f(a + a_1) = f(a) + D_a f(a_1) + \frac{1}{2!} D_a^2 f(a_1, a_1) + \ldots + \frac{1}{k!} D_a^k f(\underbrace{a_1, \ldots, a_1}_{k}) \\ + \frac{1}{k!} \int_0^1 \Big(\alpha \mapsto (1-\alpha)^k\, D_{a+\alpha a_1}^{k+1} f(\underbrace{a_1, \ldots, a_1, a_1}_{k+1})\Big). \end{aligned} \tag{5.6}$$

Proof. Let $a_1 \in U - a$, then $a, a + a_1 \in U$ and since U is convex and open, there exists an $\epsilon \in]0, \infty[$ such that $a + \alpha\, a_1 = \alpha\, a + (1-\alpha)((a + a_1) - a) \in U$ for all $\alpha \in S :=]-\epsilon, 1 + \epsilon[\subseteq \mathbb{R}$.

Let $k \in \mathbb{N}$, suppose $f \in C^k(U, B)$, and let $0 \leq l < k$. The map $g : S \to A : \alpha \mapsto a + \alpha\, a_1$ is C^1 with derivative $D_\alpha g(\beta) = \beta\, a_1$ for all $\alpha \in S$. Suppose $l = 0$, choose $h_0 : S \to B$, $h_0(\alpha) := f(a + \alpha\, a_1) = (f \circ g)(\alpha)$. By Theorem (5.1.8) we have (as $l < k$) $h_0 \in C^1(U, B)$ with derivative

$$D_\alpha h_0(\beta) = D_{g(\alpha)} f(D_\alpha g(\beta)) = D_{a + \alpha\, a_1} f(\beta\, a_1).$$

Suppose $0 < l < k$, choose $h_l : S \to B$, $h_l(\alpha) := D^l_{a+\alpha\, a_1} f(a_1, \ldots, a_1) = D^l_{g(\alpha)} f(a_1, \ldots, a_1)$. Then by Theorem (5.1.8), $h_l \in C^1(U, B)$ and has derivative

$$\begin{aligned} D_\alpha h_l(\beta) &= D^{l+1}_{g(\alpha)} f(a_1, \ldots, a_1, D_\alpha g(\beta)) \\ &= D^{l+1}_{a+\alpha\, a_1} f(a_1, \ldots, a_1, \beta a_1). \end{aligned}$$

Suppose $k = 0$ and $f \in C^{0+1}(U, B)$. Then by Theorem (5.3.8) we have

$$\begin{aligned} &\frac{1}{0!} \int_0^1 (\alpha \mapsto (1-\alpha)^0 \, D^1_{a+\alpha a_1} f(a_1)) \\ &= \int_0^1 (\alpha \mapsto D_\alpha h_0(1)) \\ &= h_0(1) - h_0(0) \\ &= f(a + a_1) - f(a). \end{aligned}$$

Therefore

$$f(a + a_1) = f(a) + \frac{1}{0!} \int_0^1 (\alpha \mapsto (1-\alpha)^0 \, D^1_{a+\alpha a_1} f(a_1))$$

and Equation (5.6) holds for $k = 0$.

Now suppose Equation (5.6) is true for $k \in \mathbb{N}$ and that $f \in C^{(k+1)+1}(U, B)$. Choose $i : S \to \mathbb{K}$, $i(\alpha) := \frac{-1}{(k+1)!} (1-\alpha)^{k+1}$, then $i \in C^1(S, \mathbb{K})$ with $D_\alpha i(1) = \frac{1}{k!} (1-\alpha)^k$, so using Lemma (5.3.10) we find

$$\begin{aligned} &\frac{1}{k!} \int_0^1 \Big(\alpha \mapsto (1-\alpha)^k \, D^{k+1}_{a+\alpha a_1} f(a_1, \ldots, a_1)\Big) \\ &= \int_0^1 \Big(\alpha \mapsto D_\alpha i(1) \, h_{k+1}(\alpha)\Big) \\ &= i(1)\, h_{k+1}(1) - i(0)\, h_{k+1}(0) - \int_0^1 \Big(\alpha \mapsto i(\alpha)\, D_\alpha h_{k+1}(1)\Big) \\ &= 0 - \frac{-1}{(k+1)!} D^{k+1}_a f(a_1, \ldots, a_1) \\ &\quad - \frac{-1}{(k+1)!} \int_0^1 \Big(\alpha \mapsto (1-\alpha)^{k+1} \, D^{k+2}_{a+\alpha a_1} f(a_1, \ldots, a_1, a_1)\Big). \end{aligned}$$

Since Equation (5.6) was true for k, we find

$$\begin{aligned} f(a + a_1) &= f(a) + D_a f(a_1) + \frac{1}{2!} D^2_a f(a_1, a_1) + \ldots + \frac{1}{k!} D^k_a f(a_1, \ldots, a_1) \\ &\quad + \frac{1}{k!} \int_0^1 \Big(\alpha \mapsto (1-\alpha)^k \, D^{k+1}_{a+\alpha a_1} f(a_1, \ldots, a_1)\Big) \\ &= f(a) + D_a f(a_1) + \frac{1}{2!} D^2_a f(a_1, a_1) + \ldots + \frac{1}{k!} D^k_a f(a_1, \ldots, a_1) \\ &\quad + \frac{1}{(k+1)!} D^{k+1}_a f(a_1, \ldots, a_1, a_1) \\ &\quad + \frac{1}{(k+1)!} \int_0^1 \Big(\alpha \mapsto (1-\alpha)^{k+1} \, D^{k+2}_{a+\alpha a_1} f(a_1, \ldots, a_1, a_1, a_1)\Big). \end{aligned}$$

Hence Equation (5.6) holds for $k+1$. So with induction, Equation (5.6) holds for all $k \in \mathbb{N}$. $\square$

Corollary 5.3.12: Taylor approximation
Let A, B Vs $/\mathbb{K}$ T2 LC UC, $U \subseteq A$ open abc, $k \in \mathbb{N}$.

- Suppose $f \in C^{k+1}(U, B)$, then for any $a_1 \in U - a$ we have

$$\lim_{\alpha \to 0} \left[\frac{1}{\alpha^{k+1}} \left(f(a + \alpha a_1) - \sum_{l=0}^{k} \frac{\alpha^l}{l!} D_a^l f(\underbrace{a_1, \ldots, a_1}_{l}) \right) \right] = \frac{1}{(k+1)!} D_a^{k+1} f(\underbrace{a_1, \ldots, a_1}_{k+1}).$$

  In particular,

$$f(a + \alpha\, a_1) = f(a) + \frac{\alpha^1}{1!} D_a^1 f(a_1) + \ldots + \frac{\alpha^k}{k!} D_a^k f(a_1, \ldots, a_1) + \mathcal{O}(\alpha^{k+1}).$$

- Suppose $f \in C^k(U, B)$ and let $\|\cdot\| : A \to \mathbb{R}$ C be a seminorm. Suppose there exists an $\epsilon \in]0, \infty[$ such that for all $a_1 \in U - a$ and all $\alpha \in [0, 1]$

$$\left\| D^k_{a+\alpha a_1} f(a_1, \ldots, a_1) - D^k_a f(a_1, \ldots, a_1) \right\| \leq \epsilon^k,$$

  then for any $a_1 \in U - a$ we have

$$\left\| f(a + a_1) - \sum_{l=0}^{k} \frac{1}{l!} D_a^l f(\underbrace{a_1, \ldots, a_1}_{l}) \right\| \leq \frac{\epsilon^k}{k!}.$$

Proof. • Suppose $f \in C^{k+1}(U, B)$ and let $a_1 \in U - a$, $\alpha \in B_{\mathbb{K}}(0, 1) \setminus \{0\}$, then from Equation (5.6) we find

$$\begin{aligned}
&\frac{1}{\alpha^{k+1}} \left(f(a + \alpha\, a_1) - \sum_{l=0}^{k} \frac{\alpha^l}{l!} D_a^l f(\underbrace{a_1, \ldots, a_1}_{l}) \right) \\
&= \frac{1}{\alpha^{k+1}} \left(f(a + \alpha\, a_1) - f(a) - \ldots - \frac{1}{k!} D_a^k f(\alpha\, a_1, \ldots, \alpha\, a_1) \right) \\
&= \frac{1}{\alpha^{k+1}} \left(\frac{1}{k!} \int_0^1 \left(\beta \mapsto (1 - \beta)^k\, D^{k+1}_{a+\beta\,\alpha\, a_1} f(\alpha\, a_1, \ldots, \alpha\, a_1) \right) \right) \\
&= \frac{\alpha^{k+1}}{k! \alpha^{k+1}} \int_0^1 \left(\beta \mapsto (1 - \beta)^k\, D^{k+1}_{a+\beta\,\alpha\, a_1} f(a_1, \ldots, a_1) \right).
\end{aligned}$$

Hence

$$\begin{aligned}
&\lim_{\alpha\to 0}\frac{1}{\alpha^{k+1}}\left(f(a+\alpha\, a_1)-\sum_{l=0}^{k}\frac{\alpha^l}{l!}D_a^l f(\underbrace{a_1,\dots,a_1}_{l})\right)\\
&=\lim_{\alpha\to 0}\left(\frac{1}{k!}\int_0^1\Big(\beta\mapsto(1-\beta)^k\,D^{k+1}_{a+\beta\,\alpha\,a_1}f(a_1,\dots,a_1)\Big)\right)\\
&=\frac{1}{k!}\int_0^1\Big(\beta\mapsto(1-\beta)^k\,D^{k+1}_{a+\beta\,0\,a_1}f(a_1,\dots,a_1)\Big)\\
&=\frac{1}{(k+1)!}D_a^{k+1}f(a_1,\dots,a_1),
\end{aligned}$$

where in the beforelast step we used the fact that the map $\mathbb{K}\times\mathbb{R}\to B:(\alpha,\beta)\mapsto(1-\beta)^k\,D^{k+1}_{a+\beta\alpha a_1}f(a_1,\dots,a_1)$ ⓒ together with Lemma (5.3.7).

- Suppose $f\in C^k(U,B)$, then we find with Equation (5.6) for any $a_1\in U-a$ that

$$\begin{aligned}
f(a+a_1)&=f(a)+D_af(a_1)+\frac{1}{2!}D_a^2f(a_1,a_1)+\dots+\frac{1}{(k-1)!}D_a^{k-1}f(a_1,\dots,a_1)\\
&\quad+\frac{1}{(k-1)!}\int_0^1\Big(\alpha\mapsto(1-\alpha)^{k-1}\,D^k_{a+\alpha\,a_1}f(a_1,\dots,a_1,a_1)\Big)\\
&\quad+\frac{1}{k!}D_a^kf(a_1,\dots,a_1,a_1)\\
&\quad-\left(\int_0^1\left(\alpha\mapsto\frac{(1-\alpha)^{k-1}}{(k-1)!}\right)\right)D_a^kf(a_1,\dots,a_1,a_1)\\
&=\sum_{l=0}^{k}\frac{1}{l!}D_a^lf(\underbrace{a_1,\dots,a_1}_{l})\\
&\quad+\frac{1}{(k-1)!}\int_0^1\Big(\alpha\mapsto(1-\alpha)^{k-1}\Big(D^k_{a+\alpha\,a_1}f(a_1,\dots,a_1)\\
&\quad-D_a^kf(a_1,\dots,a_1)\Big)\Big).
\end{aligned}$$

Now using Theorem (5.3.5) and the fact that $\|\cdot\|$ is ⓒ and a seminorm, we find

$$\begin{aligned}
&\left\|\frac{1}{(k-1)!}\int_0^1\Big(\alpha\mapsto(1-\alpha)^{k-1}\Big(D^k_{a+\alpha\,a_1}f(a_1,\dots,a_1)-D_a^kf(a_1,\dots,a_1)\Big)\Big)\right\|\\
&\le\frac{1}{(k-1)!}\int_0^1\Big(\alpha\mapsto(1-\alpha)^{k-1}\left\|D^k_{a+\alpha\,a_1}f(a_1,\dots,a_1)-D_a^kf(a_1,\dots,a_1)\right\|\Big)\\
&\le\frac{1}{(k-1)!}\int_0^1\Big(\alpha\mapsto(1-\alpha)^{k-1}\,\epsilon^k\Big)\\
&=\frac{\epsilon^k}{k!}
\end{aligned}$$

from which the estimate follows.

□

Finally, Theorem (5.3.8) permits us to show that the derivative introduced in Definition (5.1.1) is equivalent to the Gâteaux derivative, as introduced in [Ham1982], if the involved function is continuously differentiable.

Theorem 5.3.13: Compatibility with Gâteaux derivative
Let A, B Vs $/\mathbb{K}$ T2 LC UC, $U \subseteq A$ open abc, $f : U \to B$.

Then $f \in C^1(U, B)$ if and only if there exists a map $g : U \times U \times A \to B$ c (compare with Lemma 3.3.1 of [Ham1982]), such that for all $a_1, a_2 \in U$, $A \to B : a_3 \mapsto g(a_1, a_2, a_3)$ l $/\mathbb{K}$ and

$$f(a_2) - f(a_1) = g(a_1, a_2, a_2 - a_1).$$

If this is the case, then for all $a \in U$, $a_1 \in A$,

$$D_a f(a_1) = g(a, a, a_1).$$

Proof. We follow [Ham1982].

Suppose such a map g exists. Fix $a \in U$, $a_1 \in A$. Let $a_2 \in A$ and $\alpha \in \mathbb{K}$ small enough but nonzero. Then

$$\begin{aligned} \frac{1}{\alpha}\Big(f(a + \alpha\, a_2) - f(a)\Big) &= \frac{1}{\alpha}\, g(a, a + \alpha\, a_2, \alpha\, a_2) \\ &= g(a, a + \alpha\, a_2, a_2) \end{aligned}$$

so as g c,

$$\lim_{(\alpha, a_2) \to (0, a_1)} \frac{1}{\alpha}\Big(f(a + \alpha\, a_2) - f(a)\Big) = g(a, a + 0\, a_1, a_1).$$

Since the map $A \to B : a_2 \mapsto g(a, a, a_2)$ c l $/\mathbb{K}$ we find with Lemma (5.1.4) that f d a and $D_a f(a_1) = g(a, a, a_1)$. Furthermore, g c by assumption, so $(a, a_1) \mapsto D_a f(a_1) = g(a, a, a_1)$ c, and hence $f \in C^1(U, B)$.

Suppose conversely that $f \in C^1(U, B)$. Then we can define $g : U \times U \times A \to B$ by

$$g(a_1, a_2, a_3) := \int_0^1 \Big(\alpha \mapsto D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_3)\Big).$$

Then by the fact that $(a, a_1) \mapsto D_a f(a_1)$ c and Lemma (5.3.7) we find that g c. By Theorem (5.3.5) and the fact that $D_a f$ l we also find that for any $a_1, a_2 \in U$, $a_3 \mapsto g(a_1, a_2, a_3)$ l.

Furthermore, by Theorem (5.3.8), for any $a_1, a_2 \in U$

$$\begin{aligned} f(a_2) - f(a_1) &= f(1\, a_2 - (1-1)\, a_1) - f(0\, a_2 - (1-0)\, a_1) \\ &= \int_0^1 \Big(\alpha \mapsto D_{\alpha\, a_2 - (1-\alpha)\, a_1} f(a_2 - a_1)\Big) \\ &= g(a_1, a_2, a_2 - a_1). \end{aligned}$$

□

We are now going to prove a result that will be necessary for the treatment of geodesics in Section 6.4. This statement originates from the study of classical mechanics, see [Dui2006].

Theorem 5.3.14: Euler-Lagrange variational formula ($\mathrm{La}(f,g)$)
Let A, B TVS $/\mathbb{K}$ T2 LC UC, $U \subseteq A \times A$ open.

Define for $f \in C^2(U,B)$ and any path $g : S \to A$ with $S \subseteq \mathbb{R}$ an open interval, $g \in C^2(S,A)$, and $(g(\alpha), g'(\alpha)) \in U$ for all $\alpha \in S$, the *Lagrange map of g with respect to f* as $\mathrm{La}(f,g) : S \times A \to B$, given by the family of linear maps

$$\mathrm{La}(f,g)(\alpha,a) := \frac{\partial}{\partial\alpha}\Big(\Big(\frac{\partial f(g(\alpha),a_2)}{\partial a_2}\Big|_{a_2=g'(\alpha)}\Big)(a)\Big) - \Big(\frac{\partial f(a_1,g'(\alpha))}{\partial a_1}\Big|_{a_1=g(\alpha)}\Big)(a).$$

Let C TVS $/\mathbb{K}$ T2 LC UC, $W \subseteq C$ open and a family of paths $g : W \times S \to A : (c,\alpha) \mapsto g_c(\alpha)$, $g \in C^2(W \times S, A)$, $S \subseteq \mathbb{R}$ an open interval, such that $(g_c(\alpha), g_c'(\alpha)) \in U$ for all $c \in W$, $\alpha \in S$.

Then we have the *Euler-Langrage variational formula*: for any $\gamma, \delta \in S$, $c \in W$, and $c_1 \in C$ we have

$$\begin{aligned}\frac{\partial}{\partial c}\Big(\int_\gamma^\delta \Big(\alpha \mapsto f(g_c(\alpha), g_c'(\alpha))\Big)\Big)(c_1) &= -\int_\gamma^\delta \Big(\alpha \mapsto \mathrm{La}(f,g_c)\Big(\alpha, \frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\Big)\\ + \frac{\partial f(g_c(\delta),a_2)}{\partial a_2}\Big|_{a_2=g_c'(\delta)}\Big(\frac{\partial g_c(\delta)}{\partial c}(c_1)\Big) &- \frac{\partial f(g_c(\gamma),a_2)}{\partial a_2}\Big|_{a_2=g_c'(\gamma)}\Big(\frac{\partial g_c(\gamma)}{\partial c}(c_1)\Big).\end{aligned} \tag{5.7}$$

Proof. We follow [Dui2006]. Note that $g_c'(\alpha) = \frac{\partial g_c(\alpha)}{\partial \alpha}(1)$. By Lemma (5.3.7), Lemma (5.1.13), and Theorem (5.1.8) we have

$$\begin{aligned}&\frac{\partial}{\partial c}\Big(\int_\gamma^\delta \Big(\alpha \mapsto f(g_c(\alpha), g_c'(\alpha))\Big)\Big)(c_1)\\ &= \int_\gamma^\delta \Big(\alpha \mapsto \frac{\partial}{\partial c}\Big(f(g_c(\alpha), g_c'(\alpha))\Big)(c_1)\Big)\\ &= \int_\gamma^\delta \Big(\alpha \mapsto \Big(\frac{\partial f(a_1, g_c'(\alpha))}{\partial a_1}\Big|_{a_1=g_c(\alpha)}\Big)\Big(\frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\\ &\qquad + \Big(\frac{\partial f(g_c(\alpha),a_2)}{\partial a_2}\Big|_{a_2=g_c'(\alpha)}\Big)\Big(\frac{\partial^2 g_c(\alpha)}{\partial c\,\partial\alpha}(1,c_1)\Big)\Big).\end{aligned}$$

Note that with Theorem (5.1.8) and Equation (5.4)

$$\begin{aligned}&\frac{\partial}{\partial\alpha}\Big(\Big(\frac{\partial f(g_c(\alpha),a_2)}{\partial a_2}\Big|_{a_2=g_c'(\alpha)}\Big)\Big(\frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\Big)\\ &= \mathrm{La}(f,g_c)\Big(\alpha, \frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big) + \frac{\partial f(a_1, g_c'(\alpha))}{\partial a_1}\Big|_{a_1=g_c(\alpha)}\Big(\frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\\ &\quad + \Big(\frac{\partial f(g_c(\alpha),a_2)}{\partial a_2}\Big|_{a_2=g_c'(\alpha)}\Big)\Big(\frac{\partial^2 g_c(\alpha)}{\partial\alpha\,\partial c}(c_1,1)\Big).\end{aligned}$$

So with Theorem (5.1.16) we obtain from our first expression that

$$\begin{aligned}
&\frac{\partial}{\partial c}\Big(\int_\gamma^\delta \Big(\alpha \mapsto f(g_c(\alpha), g_c'(\alpha))\Big)\Big)(c_1)\\
&= \int_\gamma^\delta \Big(\alpha \mapsto -\operatorname{La}(f,g)\Big(\alpha, \frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\\
&\qquad + \frac{\partial}{\partial \alpha}\Big(\Big(\frac{\partial f(g_c(\alpha), a_2)}{\partial a_2}\Big|_{a_2 = g_c'(\alpha)}\Big)\Big(\frac{\partial g_c(\alpha)}{\partial c}(c_1)\Big)\Big)\Big)
\end{aligned}$$

which yields Equation (5.7) via Theorem (5.3.5) and Theorem (5.3.8). □

5.4 Fréchet spaces

Definition 5.4.1: Fréchet space (Fr)

Let A be a set.

Then we call A a *Fréchet space* (denoted by A Fr $/\mathbb{K}$) if A Vs $/\mathbb{K}$ T2 LC UC for which the collection of seminorms giving rise to local convexity is countable.

Using Lemma (4.5.10) we see that for any A Fr $/\mathbb{K}$, A FS $/\mathbb{K}$.

Therefore almost all theory derived in the previous sections is valid for Fréchet spaces; we summarise these results in Theorem (5.4.2) for convenience.

Theorem 5.4.2

Let A, B Fr $/\mathbb{K}$ with seminorms $\{\|\cdot\|_i | i \in \mathbb{N}\}$, $\{\|\cdot\|_j' | j \in \mathbb{N}\}$ respectively.

- Let $f : A \to B$ be a map, $a \in A$, $b \in B$.

  Then the following are equivalent:

  - $\lim_{x \to a} f(x) = b$,
  - for all $j \in \mathbb{N}$ and $\epsilon \in]0, \infty[$ there exist $i_1, \ldots, i_k \in \mathbb{N}$ and a $\delta \in]0, \infty[$ such that for all $a_1 \in A$ with $\|a_1 - a\|_{i_1} < \delta$, $\ldots$, $\|a_1 - a\|_{i_k} < \delta$ we have $\|f(a_1) - b\|_j' < \epsilon$,
  - for all sequences $x : \mathbb{N} \to A$ with $\lim_{k \to \infty} x_k = a$ we have for all $j \in \mathbb{N}$ that $\lim_{k \to \infty} \|f(x_k) - b\|_j' = 0$.

- Let $a \in A$. Then the following are equivalent:

  - $a = 0$,
  - for all $i \in \mathbb{N}$, $\|a\|_i = 0$,
  - for all $f \in A'$, $f(a) = 0$.

- Suppose $B \leq A$. Then for any $f \in B'$ there exists a $g \in A'$ with $g|_B = f$.

- We have that $A \simeq (A')'$ are Vs-isomorphic.

- Let $f : A \to B$ l $/\mathbb{K}$. Then the following are equivalent:

  - f c,
  - $\operatorname{graph}(f) \subseteq A \times B$ is closed,

  – for all $j \in \mathbb{N}$ there exists an $\alpha \in]0, \infty[$ and $i_1, \ldots, i_k \in \mathbb{N}$ such that for all $a \in A$ we have

$$\|f(a)\|'_j \leq \alpha\Big(\|a\|_{i_1} + \ldots + \|a\|_{i_k}\Big).$$

- Let $f : A \to B$ c l $/\mathbb{K}$.

  Then

  – f is surjective if and only if f is open,

  – f is bijective if and only if $f^{-1} : B \to A$ c l $/\mathbb{K}$ if and only if f is a TVs-isomorphism.

Proof. • Use Lemma (4.5.4) for the first equivalence, together with the fact that the topology on A and B is the initial topology of their respective seminorms, Lemma (2.1.19). By Lemma (4.5.10), A FS and hence A d(.,.). Therefore by Theorem (2.5.10) and Lemma (2.3.4) we obtain equivalence with the third item.

- First of all if $a = 0$, then by definition $\|a\|_i = \|0\,a\|_i = 0\,\|a\|_i = 0$ and $f(a) = f(0\,a) = 0\,f(a) = 0$ for any $i \in I$, $f \in A'$. Conversely use Theorem (4.5.14) and Lemma (4.5.6).

- This is Theorem (4.5.14).

- By Theorem (4.5.15) we have that $f : A \to (A')' : a \mapsto (g \mapsto g(a))$ c l $/\mathbb{K}$ and bijective. Hence by Corollary (4.4.6) and the fact that A FS we see that f is a TVs-isomorphism.

- Use Theorem (4.4.5) and Lemma (4.5.5).

- Use Theorem (4.4.5) and Corollary (4.4.6).

□

Despite Theorem (5.4.2), Section 5.1, and Section 5.3 there are still quite a few results which are not valid in general Fréchet spaces, among which the inverse function theorem and existence and uniqueness of solutions of ordinary differential equations.

An extensive treatment of a version of the inverse function theorem that can be applied in a broader context (the Nash-Moser inverse function theorem) is given in [Ham1982]. Here we will only include two counterexamples from this article.

Example 5.4.3: Fréchet spaces and the inverse function theorem

Let

$$A := \{f : \mathbb{C} \to \mathbb{C} \mid f \text{ holomorphic on } \mathbb{C}\}$$

with topology induced by the seminorms

$$\|f\|_{k,B} := \sup\{|f^{(l)}(x)| \in \mathbb{R} | 0 \leq l \leq k, x \in B\}$$

for all $k \in \mathbb{N}$ and $B \subseteq \mathbb{C}$ Cpt (here $f^{(l)}$ denotes the l-th derivative of f). As $\mathbb{Q}^2 \subseteq \mathbb{C}$ is countable and dense we can make a countable selection of seminorms which makes A Fr $/\mathbb{C}$.

Consider the exponential map $F : A \to A$ given by

$$F(f)(x) := e^{f(x)}$$

for all $f \in A$, $x \in \mathbb{C}$.

Then $F \in C^\infty(A, A)$ with derivative given for all $k \in \mathbb{N}$ and $f, f_1, \ldots, f_k \in A$ by

$$\Big(D_f^k F(f_1, \ldots, f_k)\Big)(x) = e^{f(x)}\, f_1(x)\, \ldots\, f_k(x).$$

In particular $D_f F$ is bijective for all $f \in A$ with inverse given by

$$[D_f F]^{-1} : A \to A : f_1 \mapsto \Big(x \mapsto e^{-f(x)}\, f_1(x)\Big).$$

It is clear that the family of inverses

$$[DF]^{-1} : A \times A \to A : (f, f_1) \mapsto [D_f F]^{-1}(f_1)$$

is also C^∞ as a function $A \times A \to A$.

Now let

$$B := \{f \in A \mid \forall x \in \mathbb{C} : f(x) \neq 0\} \subseteq A$$

then $F(A) \subseteq B$ by definition, since $|F(f)(x)| = |e^{f(x)}| > 0$ for all $f \in A$, $x \in \mathbb{C}$.

Suppose that there is a nonempty open $U \subseteq A$ for which $U \subseteq B$.

Let $f \in U$, then because $f \in A$ is holomorphic there exists a sequence of polynomials $\mathbb{N} \to A : k \mapsto p_k$, for example

$$p_k(x) := f(0) + \frac{f^{(1)}(0)}{1!}\, x^1 + \ldots + \frac{f^{(k)}(0)}{k!}\, x^k,$$

for which $\lim_{k\to\infty} p_k = f$. Since U is an open neighbourhood of f in A, there exists a $k \in \mathbb{N}$ such that $p_l \in U$ for all $l \geq k$. In particular $p_k \in U \subseteq B$ so p_k has no zeroes; this is in contradiction with the fundamental theorem of calculus: p_k has $k > 0$ zeroes.

Therefore $B \supseteq F(A)$ cannot contain a nonempty open set and hence F does not have a differentiable inverse defined on any open $U \subseteq A$, even though $F \in C^\infty(A, A)$, $D_f F$ is invertible for all $f \in A$, and the family of inverses $[DF]^{-1} \in C^\infty(A \times A, A)$.

Example 5.4.4: Fréchet spaces and ordinary differential equations
Let

$$A := C^\infty([-1, 1], \mathbb{R})$$

with topology induced by the seminorms

$$\|f\|_k := \sup\{|f^{(l)}(x)| \in \mathbb{R} \mid 0 \leq l \leq k, x \in [-1, 1]\}$$

which make A Fr $/\mathbb{R}$.

Let $F : A \to A$ be the map

$$F(f)(x) := f'(x).$$

Then $F \in C^\infty(A, A)$ because F c l $/\mathbb{R}$.

Now consider the ordinary differential equation for fixed $f_0 \in A$

$$\gamma'(t) = F(\gamma(t)), \qquad \gamma(0) = f_0$$

where $\gamma : I \to A$ for an open interval $0 \in I \subseteq \mathbb{R}$ is the sought solution.

Writing $\gamma(t, x) = \gamma(t)(x)$ we see that the differential equation can be written as

$$\frac{\partial \gamma}{\partial t}(t, x) = \frac{\partial \gamma}{\partial x}(t, x), \qquad \gamma(0, x) = f_0(x)$$

for all $t \in I$, $x \in [-1, 1]$. This admits solutions $\gamma(t, x) = f(t + x)$ where $f : I + [-1, 1] \to \mathbb{R}$ is a smooth function satisfying $f|_{[-1,1]} = f_0$.

Hence the solution is by no means unique (e.g. take the function $f_0(x) = e^{\frac{-1}{(x-1)^2}}$ for $x \in]-1, 1[$, $f_0(\pm 1) = 0$ then $f_0^{(k)}(\pm 1) = 0$ for all $k \in \mathbb{N}$, so we can find a myriad of different extensions f of f_0).

On the other hand, if we would take

$$A := \{f : \mathbb{R} \to \mathbb{R} \mid f|_{[-1,1]} \in C^\infty([-1, 1]), \forall x \le -1 : f(x) = 0\},$$

then the differential equation does not have any solution if $f_0(x) \neq 0$ for $x \in]-1, 1[$ (e.g. again consider $f_0(x) = e^{\frac{-1}{(x-1)^2}} > 0$ for $x \in]-1, 1[$ and $f_0(x) = 0$ for $x \notin]-1, 1[$).

5.5 Banach spaces

Definition 5.5.1: Banach (Ba)

Let A be a set.

Then we call A a *Banach space* (denoted by A Ba $/\mathbb{K}$) if A $\|\cdot\|$ $/\mathbb{K}$ (recall Definition (4.2.3) and Definition (4.2.4)) which is complete as a metric space.

Lemma 5.5.2

Let A Vs $/\mathbb{K}$.

Suppose A Ba $/\mathbb{K}$, then A Fr $/\mathbb{K}$.

Conversely, if A Fr $/\mathbb{K}$ and the collection of seminorms giving rise to local convexity is finite, then A Ba $/K$.

Proof. By Lemma (4.5.7) we see that A $\|\cdot\|$ if and only if A Vs T2 LC with a finite number of seminorms. Suppose A Ba $/\mathbb{K}$, then A is complete by definition and since A only has a single norm, this implies that A Vs T2 LC UC with a single seminorm and hence A Fr. Suppose conversely that A Fr with a finite number of seminorms, then A $\|\cdot\|$ and by Lemma (4.5.10), A is complete, so A Ba. $\square$

In particular all results in Theorem (5.4.2) hold for Ba.

Lemma 5.5.3

Let $k \in \mathbb{N}$, $A_1, \ldots, A_k, B$ Ba $/K$, $f : A_1 \times \ldots \times A_k \to B$ k-1 $/\mathbb{K}$.

Then f C if and only if there exists an $\alpha \in]0, \infty[$ such that for all $a_1 \in A_1$, $\ldots$, $a_k \in A_k$ we have

$$\|f(a_1, \ldots, a_k)\|_B \le \alpha \|a_1\|_{A_1} \ldots \|a_k\|_{A_k}.$$

Furthermore if A Ba $/\mathbb{K}$, $U \subseteq A$ open, and $f : U \times A_1 \times \ldots \times A_k \to B$ C a k-linear family, then for all $a \in U$ there exists a $\delta \in]0, \infty[$ and $\alpha \in]0, \infty[$ such that for all $a' \in B_A(a, \delta) \subseteq U$, $a_1 \in A_1$, $\ldots$, $a_k \in A_k$ we have

$$\|f_{a'}(a_1, \ldots, a_k)\|_B \le \alpha \|a_1\|_{A_1} \ldots \|a_k\|_{A_k}. \tag{5.8}$$

Proof. Suppose that the estimate holds for all $a_1 \in A_1$, $\dots$, $a_k \in A_k$. Let $\epsilon \in]0,\infty[$ be given, then pick $\delta = \sqrt[k]{\frac{\epsilon}{\alpha}} \in]0,\infty[$ to obtain for all $a_1 \in B_{A_1}(0,\delta)$, $\dots$, $a_k \in B_{A_k}(0,\delta)$ that $\|f(a_1,\dots,a_k)\|_B \le \alpha\,\|a_1\|_{A_1}\dots\|a_k\|_{A_k} < \alpha\,\delta^k = \epsilon$. Hence $\lim_{(a_1,\dots,a_k)\to(0,\dots,0)} f(a_1,\dots,a_k) = 0$ and since f k-l we have f c. ◻

Now let A Ba, $U \subseteq A$ open and $f : U \times A_1 \times \dots \times A_k \to B$ c a k-linear family. Let $a \in U$. Then $\lim_{(a',a_1,\dots,a_k)\to(a,0,\dots,0)} f_{a'}(a_1,\dots,a_k) = f_a(0,\dots,0) = 0$, so for $1 \in]0,\infty[$ there exists a $\delta \in]0,\infty[$ such that $B_A(a,\delta) \subseteq U$ and for all $a' \in B_A(a,\delta)$, $a_1 \in B_{A_1}(0,2\,\delta)$, $\dots$, $a_k \in B_{A_k}(0,2\,\delta)$ we have $f(a_1,\dots,a_k) \in B_B(0,1)$. Let $a' \in B_A(a,\delta)$, $a_1 \in A_1$, $\dots$, $a_k \in A_k$, then if for $1 \le l \le k$ some $\|a_l\|_{A_l} = 0$, $a_l = 0$, so $\|f_{a'}(a_1,\dots,a_k)\|_B = \|0\|_B = 0$ and Equation (5.8) holds. Otherwise note that for $1 \le l \le k$, $\|\frac{\delta}{\|a_l\|_{A_l}}\,a_l\|_{A_l} = \delta\frac{\|a_l\|_{A_l}}{\|a_l\|_{A_l}} < 2\,\delta$, so $\frac{\delta}{\|a_l\|_{A_l}}\,a_l \in B_{A_l}(0,2\,\delta)$ and hence $\frac{\delta}{\|a_1\|_{A_1}}\cdots\frac{\delta}{\|a_k\|_{A_k}}\,\|f_{a'}(a_1,\dots,a_k)\|_B = \|f_{a'}(\frac{\delta}{\|a_1\|_{A_1}}\,a_1,\dots,\frac{\delta}{\|a_k\|_{A_k}}\,a_k)\|_B < 1$, so $\|f_{a'}(a_1,\dots,a_k)\|_B \le \frac{1}{\delta^k}\,\|a_1\|_{A_1}\dots\|a_k\|_{A_k}$. Hence if we pick $\alpha = \frac{1}{\delta^k} \in]0,\infty[$, then we see that Equation (5.8) holds. ◻

Definition 5.5.4: Space of all continuous linear maps ($L(A,B)$)
Let A, B Ba $/\mathbb{K}$.

Define *the space of all continuous linear maps between A and B* (denoted by $L(A,B)$) by

$$L(A,B) := \{f : A \to B \mid f \text{ c l } /\mathbb{K}\ \}$$

together with $0(a) := 0$, $(f+g)(a) := f(a) + g(a)$, $(\alpha\,f)(a) := \alpha\,f(a)$ and the norm $\|\cdot\|_\infty : L(A,B) \to \mathbb{R}$ defined by

$$\|f\|_\infty := \sup\{\|f(a)\|_B/\|a\|_A \in \mathbb{R} \mid a \in A\setminus\{0\}\}.$$

Note that this definition implies that for all $f \in L(A,B)$ and $a \in A$ we have

$$\|f(a)\|_B \le \|f\|_\infty\,\|a\|_A$$

and that for all $f \in L(A,B)$, $\|f\|_\infty < \infty$ exists by Lemma (5.5.3).

Lemma 5.5.5
Let A, B Ba $/\mathbb{K}$, $U \subseteq A$ open abc, $f \in C^1(U,B)$.

If there exists a $\delta \in]0,\infty[$ such that $\|D_af\|_\infty \le \delta$ for all $a \in U$, then f δ-c: for all $a_1, a_2 \in U$ we have that

$$\|f(a_2) - f(a_1)\|_B \le \delta\,\|a_2 - a_1\|_A.$$

Proof. Let $a_1, a_2 \in U$. Since U is convex by assumption, $\alpha\,a_2 + (1-\alpha)\,a_1 \in U$

for all $\alpha \in [0,1]$. Hence by Theorem (5.3.5) and Theorem (5.3.8)

$$\begin{aligned}
\|f(a_2) - f(a_1)\|_B &= \|f(1\, a_2 + (1-1)\, a_1) - f(0\, a_2 + (1-0)\, a_1)\|_B \\
&= \left\| \int_0^1 \left(\alpha \mapsto D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_2 - a_1) \right) \right\|_B \\
&\leq \int_0^1 \left(\alpha \mapsto \|D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_2 - a_1)\|_B \right) \\
&\leq \int_0^1 \left(\alpha \mapsto \|D_{\alpha\, a_2 + (1-\alpha)\, a_1} f\|_\infty \, \|a_2 - a_1\|_A \right) \\
&\leq \int_0^1 \left(\alpha \mapsto \delta \, \|a_2 - a_1\|_A \right) \\
&= \delta \, \|a_2 - a_1\|_A.
\end{aligned}$$

□

Theorem 5.5.6: Properties of $L(A,B)$
Let A, B Ba $/\mathbb{K}$.
Then

- $L(A,B)$ Ba $/\mathbb{K}$,
- for C Ba $/\mathbb{K}$, $f \in L(A,B)$, $g \in L(B,C)$, we have
$$\|g \circ f\|_\infty \leq \|g\|_\infty \, \|f\|_\infty,$$
- for any $U \subseteq A$ open and map $f : U \to L(B,C)$ C, the induced map
$$g : U \times B \to C : (a,b) \mapsto f(a)(b)$$
is C,
- for any $f \in L(A,A)$ with $\|f\|_\infty < 1$, $\mathrm{id}_A - f : A \to A$ is invertible and $\|(\mathrm{id}_A - f)^{-1}\|_\infty \leq \frac{1}{1 - \|f\|_\infty}$,
- the collection of invertible maps
$$L(A,B)^* := \{ f \in L(A,B) |\ f \text{ bijective } \} \subseteq L(A,B)$$
is open and the map
$$L(A,B)^* \to L(B,A)^* : f \mapsto f^{-1}$$
is C, in particular it is a T-isomorphism.

Proof. • By Lemma (5.5.3), the supremum $\|f\|_\infty$ exists in $\mathbb{R}$ for all $f \in L(A,B)$ as $\|f(a)\|_B / \|a\|_A \leq \alpha$ for all $a \in A \setminus \{0\}$. It is also clear that $\|f\|_\infty \geq 0$ for all $f \in L(A,B)$. Suppose that $\|f\|_\infty = 0$, then necessarily $\|f(a)\|_B = 0$ for all $a \in A$, so $f(a) = 0$ for all $a \in A$ and hence $f = 0$. As for all $a \in A$, $\|(f+g)(a)\|_B = \|f(a) + g(a)\|_B \leq \|f(a)\|_B + \|g(a)\|_B$ we see that $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Furthermore for $\alpha \in \mathbb{K}$, $a \in A$, $\|(\alpha\, f)(a)\|_B = \|\alpha\, f(a)\|_B = |\alpha| \, \|f(a)\|_B$, so $\|\alpha\, f\|_\infty = |\alpha| \|f\|_\infty$. This makes $\|\cdot\|_\infty$ a norm and therefore $L(A,B)$ $\|\cdot\|$ $/\mathbb{K}$.

Let $\mathbb{N} \to L(A,B) : k \mapsto f_k$ be a sequence that is Cauchy with respect to $\|\cdot\|_\infty$. Fix $a \in A$, $a \neq 0$, and let $\epsilon \in]0,\infty[$ be given. Then because $k \mapsto f_k$ is Cauchy, there exists a $k \in \mathbb{N}$ such that for all $l,m \geq k$ we have $\|f_l - f_m\|_\infty < \frac{\epsilon}{\|a\|_A}$. Hence $\|f_l(a) - f_k(a)\|_B \leq \frac{\epsilon}{\|a\|_A} \|a\|_A = \epsilon$. This makes the sequence $k \mapsto f_k(a)$ Cauchy and therefore (B is complete) there exists an $f(a) \in B$ such that $\lim_{k\to\infty} f_k(a) = f(a)$. Use this to construct a map $f : A \to B$, define $f(0) := 0$. Since all the f_k are l and addition and scalar multiplication on A C, f l.

Let $\epsilon \in]0,\infty[$, then there exists a $k \in \mathbb{N}$ such that $\|f_l - f_m\|_\infty < \epsilon/2$ for all $l,m \geq k$. Fix $a \in A$ and $l \geq k$, then by continuity of $\|\cdot\|$ we have that $\|f_l(a) - f(a)\|_B = \lim_{m\to\infty} \|f_l(a) - f_m(a)\|_B \leq \lim_{m\to\infty} \|f_l - f_m\|_\infty \|a\|_A \leq \epsilon/2 \, \|a\|_A$. Hence $\|f_l - f\|_\infty \leq \epsilon/2 < \epsilon$ for all $l \geq k$ and therefore $\lim_{k\to\infty} f_k = f$.

Again let $\epsilon \in]0,\infty[$, then there is a $k \in \mathbb{N}$ such that $\|f_l - f\|_\infty < \epsilon/2$ for all $l \geq k$. As f_k C, $\lim_{a\to 0} f_k(a) = 0$, there exists a $\delta \in]0,1[$ such that $\|f_k(a)\|_B < \epsilon/2$ for all $a \in B_A(0,\delta)$. Now for all $a \in B_A(0,\delta)$, $\|f(a)\|_B = \|f(a) - f_k(a) + f_k(a)\|_B \leq \|f(a) - f_k(a)\|_B + \|f_k(a)\|_B \leq \|f - f_k\|_\infty \|a\|_A + \|f_k(a)\|_B < (\epsilon/2)\,1 + \epsilon/2 = \epsilon$. So $\lim_{a\to 0} f(a) = 0$, and f C. This gives us that $f \in L(A,B)$.

Therefore, for any Cauchy sequence $k \mapsto f_k$ in $L(A,B)$ there exists an $f \in L(A,B)$ such that $\lim_{k\to\infty} f_k = f$. So $L(A,B)$ is complete, this makes $L(A,B)$ Ba.

- Let C Ba$/\mathbb{K}$, $f \in L(A,B)$, $g \in L(B,C)$. Then for any $a \in A$ we have that $\|(g \circ f)(a)\|_C = \|g(f(a))\|_C \leq \|g\|_\infty \|f(a)\|_B \leq \|g\|_\infty \|f\|_\infty \|a\|_A$ and hence $\|(g \circ f)\|_\infty \leq \|g\|_\infty \|f\|_\infty$.

- Let $U \subseteq A$ open and $f : U \to L(B,C)$ C, define $g : U \times B \to C$ by $g(a,b) := f(a)(b)$. Then for all $a, a' \in U$, $b, b' \in B$ we have

$$\begin{aligned}
\|g(a,b) - g(a',b')\|_C &= \|f(a)(b) - f(a')(b')\|_C \\
&= \|f(a)(b) - f(a)(b') + f(a)(b') - f(a')(b')\|_C \\
&\leq \|f(a)(b) - f(a)(b')\|_C + \|f(a)(b') - f(a')(b')\|_C \\
&\leq \|f(a)\|_\infty \|b - b'\|_B + \|f(a) - f(a')\|_\infty \|b'\|_B
\end{aligned}$$

which can be made arbitrarily small by choosing b near b' and a near a' ($\lim_{a'\to a} \|f(a) - f(a')\|_\infty = 0$ as f C). Hence g C.

- Let $f \in B_{L(A,A)}(0,1)$, and define for $k \in \mathbb{N}$, $f^k := \underbrace{f \circ \ldots \circ f}_{k}$, and $f^0 := \mathrm{id}_A$. Let for $k \in \mathbb{N}$, $g_k := \sum_{l=0}^{k} f^l \in L(A,A)$. Then for $k,l \in \mathbb{N}$ we have $\|g_{k+l} - g_k\|_\infty = \|\sum_{m=1}^{l} f^{k+m}\|_\infty \leq \sum_{m=1}^{l} \|f^k \circ f^m\|_\infty \leq \sum_{m=1}^{l} \|f^k\|_\infty \|f^m\|_\infty \leq \|f\|_\infty^k \sum_{m=1}^{l} \|f\|_\infty^m \leq \|f\|_\infty^k/(1 - \|f\|_\infty)$, since $0 \leq \|f\|_\infty < 1$. Hence $\mathbb{N} \to L(A,A) : k \mapsto g_k$ is Cauchy so by completeness of $L(A,A)$ there exists a $g \in L(A,A)$ such that $g = \lim_{k\to\infty} g_k = \sum_{l=0}^{\infty} f^l$. On the other hand, for any $k \in \mathbb{N}$ we have $(\mathrm{id}_A - f) \circ g_k = g_k - f \circ g_k = \mathrm{id}_A - f^{k+1}$ so letting $k \to \infty$ we find $(\mathrm{id}_A - f) \circ g = \mathrm{id}_A - 0$ since $\|f^{k+1}\|_\infty \leq \|f\|_\infty^{k+1} \to 0$. Similarly $g \circ (\mathrm{id}_A - f) = \mathrm{id}_A$, so $\mathrm{id}_A - f$ is invertible with inverse g. Furthermore, $\|g_k\| \leq \sum_{l=0}^{k} \|f\|_\infty^k \leq \frac{1}{1 - \|f\|_\infty}$

for all $k \in \mathbb{N}$, so $\|g\| \leq \frac{1}{1-\|f\|_\infty}$. Hence $(\mathrm{id}_A - f) \in L(A,A)^*$ for all $f \in B_{L(A,A)}(0,1)$ and $\|(\mathrm{id}_A - f)^{-1}\|_\infty \leq \frac{1}{1-\|f\|_\infty}$.

- Let
$$I : L(A,B)^* \to L(B,A)^* : f \mapsto f^{-1}.$$
Let $f \in L(A,B)^*$. Note that by Corollary (4.4.6) this is equivalent to $f^{-1} \in L(B,A)^*$. Fix $\delta \in]0,1[$ and let $g \in B_{L(A,B)}(f, \delta/\|f^{-1}\|_\infty)$, then
$$\begin{aligned}\|\mathrm{id}_A - f^{-1} \circ g\|_\infty &= \|f^{-1} \circ f - f^{-1} \circ g\|_\infty \\ &= \|f^{-1} \circ (f-g)\|_\infty \\ &\leq \|f^{-1}\|_\infty \|f-g\|_\infty \\ &< \|f^{-1}\|_\infty \frac{\delta}{\|f^{-1}\|_\infty} = \delta.\end{aligned}$$
So the map $(\mathrm{id}_A - f^{-1} \circ g) \in B_{L(A,A)}(0,\delta)$ and we find that therefore $f^{-1} \circ g = \mathrm{id}_A - (\mathrm{id}_A - f^{-1} \circ g) \in L(A,A)^*$. Hence there exists a $h \in L(A,A)^*$ such that $h \circ (f^{-1} \circ g) = (f^{-1} \circ g) \circ h = \mathrm{id}_A$. Furthermore, $\|h\|_\infty \leq \frac{1}{1-\|\mathrm{id}_A - f^{-1} \circ g\|_\infty} < \frac{1}{1-\delta}$. Now $(h \circ f^{-1}) \circ g = \mathrm{id}_A$ and $\mathrm{id}_B = f \circ f^{-1} = (f \circ \mathrm{id}_A) \circ f^{-1} = (f \circ (f^{-1} \circ g \circ h)) \circ f^{-1} = g \circ (h \circ f^{-1})$, so $g \in L(A,B)^*$ with inverse $h \circ f^{-1}$. Now $\|I(g)\|_\infty = \|h \circ f^{-1}\|_\infty \leq \|h\|_\infty \|f^{-1}\|_\infty < \|f^{-1}\|_\infty/(1-\delta)$, so $I(g) \in B_{L(B,A)}(0, \|f^{-1}\|_\infty/(1-\delta))$. This means that for any $\delta \in]0,1[$ and $f \in L(A,B)^*$
$$f \in B_{L(A,B)}(f, \delta/\|f^{-1}\|_\infty) \subseteq L(A,B)^*,$$
as well as
$$I(B_{L(A,B)}(f, \delta/\|f^{-1}\|_\infty)) \subseteq B_{L(B,A)}(0, \|f^{-1}\|_\infty/(1-\delta)) \subseteq L(B,A)^*.$$
In particular we obtain the fact that $L(A,B)^*$ is open.
Let $f \in L(A,B)^*$, $\epsilon \in]0,\infty[$. Choose $\delta := \min\{\epsilon/(2\|f^{-1}\|_\infty^2), 1/(2\|f^{-1}\|_\infty)\}$, then for $g \in B_{L(A,B)}(f,\delta)$ we have by the previous item $\|g^{-1}\|_\infty < \|f^{-1}\|_\infty/(1-(1/2)) = 2\|f^{-1}\|_\infty$. Now as $f^{-1} \circ (g-f) \circ g^{-1} = (f^{-1} \circ g - \mathrm{id}_A) \circ g^{-1} = f^{-1} - g^{-1}$ we have
$$\begin{aligned}\|I(f) - I(g)\|_\infty &\leq \|f^{-1}\|_\infty \|g^{-1}\|_\infty \|g-f\|_\infty \\ &< 2\|f^{-1}\|_\infty^2 \|g-f\|_\infty \\ &< 2\|f^{-1}\|_\infty^2 \delta \leq \epsilon.\end{aligned}$$
Therefore I c.

□

Note that this implies that $L(A,A)$ together with $\|\cdot\|_\infty$ is a normed ring with multiplication defined by $(f,g) \mapsto f \circ g$.

Corollary 5.5.7: Compatibility for Banach spaces
Let A, B Ba $/\mathbb{K}$, $U \subseteq A$ open, and $f : U \to B$.

Let $a \in U$, if there exists a $g \in L(A,B)$ such that
$$\lim_{a_1 \to 0} \frac{\|f(a+a_1) - f(a) - g(a_1)\|_B}{\|a_1\|_A} = 0,$$

then f d a and $D_a f = g$.

Furthermore, if f d U and the map

$$U \to L(A, B) : a \mapsto D_a f$$

is c, then $f \in C^1(U, B)$.

Conversely, if $f \in C^2(U, B)$, then $a \mapsto D_a f$ c.

Proof. The first part of the proof follows immediately from Corollary (5.1.6).

Suppose f d U, and $U \to L(A, B) : a \mapsto D_a f$ c. Then by Theorem (5.5.6) $U \times A \to B : (a, a_1) \mapsto D_a f(a_1)$ c, so $f \in C^1(U, B)$.

Suppose $f \in C^2(U, B)$. Let $a \in U$ and $\epsilon \in]0, \infty[$. Then the 2-linear family $U \times A \times A \to B : (a', a_1, a_2) \mapsto D^2_{a'} f(a_1, a_2)$ c. So by Equation (5.8) there exists a $\delta \in]0, \infty[$ and $\alpha \in]0, \infty[$ such that for all $a' \in B_A(a, \delta) \subseteq U$, $\beta \in [0, 1]$, and $a_1, a_2 \in A$ we have (note that $\|(\beta\, a + (1-\beta)\, a') - a\|_A = (1-\beta)\|a' - a\|_A < \delta$)

$$\|D^2_{\beta\, a+(1-\beta)\, a'} f(a_1, a_2)\|_B \le \alpha\, \|a_1\|_A\, \|a_2\|_A.$$

Pick $\delta' := \min\{\frac{\epsilon}{\alpha}, \delta\} \in]0, \infty[$, then by Theorem (5.3.8) and Theorem (5.3.5) we have for any $a' \in B_A(a, \delta')$ and any $a_1 \in A$

$$\begin{aligned}
\|D_a f(a_1) - D_{a'} f(a_1)\|_B &= \|D_{1\, a+(1-1)\, a'} f(a_1) - D_{0\, a+(1-0)\, a} f(a_1)\|_B \\
&= \left\| \int_0^1 \left(\beta \mapsto D^2_{\beta\, a+(1-\beta)\, a'} f(a_1, a' - a)\right) \right\|_B \\
&\le \int_0^1 \left(\beta \mapsto \|D^2_{\beta\, a+(1-\beta)\, a'} f(a_1, a' - a)\|_B\right) \\
&\le \int_0^1 \left(\beta \mapsto \alpha\, \|a_1\|_A\, \|a' - a\|_A\right) \\
&= \alpha\, \|a_1\|_A\, \|a' - a\|_A \\
&< \alpha\, \delta' \|a_1\|_A \le \epsilon\, \|a_1\|_A.
\end{aligned}$$

Therefore, for any $a_1 \in A$, $a_1 \neq 0$,

$$\frac{\|(D_a f - D_{a'} f)(a_1)\|_B}{\|a_1\|_A} \le \epsilon$$

and hence $\|D_a f - D_{a'} f\|_\infty \le \epsilon$ for all $a' \in B_A(a, \delta')$. So $\lim_{a' \to a} D_{a'} f = D_a f$. Since this is true for all $a \in U$, $U \to L(A, B) : a \mapsto D_a f$ c. □

We are now going to prove the inverse function theorem for Banach spaces. This theorem states that if the derivative of a function at a certain point is invertible, then the function itself must also be invertible in an open neighbourhood of this point.

Theorem 5.5.8: Inverse function theorem

Let A, B Ba $/\mathbb{K}$, $U_0 \subseteq A$ open. Let $f \in C^k(U_0, B)$ such that $U_0 \to L(A, B) : a \mapsto D_a f$ c (automatically true for $k \ge 2$ by Corollary (5.5.7)).

If for a certain $a_0 \in U$ the derivative $D_{a_0} f : A \to B$ is bijective, then there exists an open neighbourhood $U \subseteq U_0$ of a_0 in A and V of $f(a_0)$ in B such that

$$f|_U : U \to V$$

is a C^k diffeomorphism.

Proof. We follow [Ham1982] and [DK2004I]. Let $f : U_0 \to B$ satisfy the hypothesis.

By Corollary (4.4.6) $D_{a_0} f : A \to A$ (being **c** **l** and bijective) is a **TVs**-isomorphism and therefore $(D_{a_0} f)^{-1}$ **c** **l**. We see therefore, by considering

$$(U_0 - a_0) \to [D_{a_0} f]^{-1}(f(U_0) - f(a_0)) : a \mapsto [D_{a_0} f]^{-1}(f(a + a_0) - f(a_0)),$$

which is C^k with derivative $[D_{a_0} f]^{-1} \circ D_{a_0} f = \mathrm{id}_A$ at 0 (Theorem (5.1.8)), that we may assume $B = A$, $U_0 \ni a_0 = 0$, $f(a_0) = 0$ and $D_{a_0} f = \mathrm{id}_A$.

The idea is that f now 'resembles' the identity mapping at 0, which we know to be a **T**-isomorphism, and we therefore consider their difference, which in turn should 'resemble' the zero mapping, let

$$g : U_0 \to A : a \mapsto a - f(a).$$

Note that $g \in C^1(U_0, A)$ with $g(0) = 0 - f(0) = 0$, $D_0 g = \mathrm{id}_A - \mathrm{id}_A = 0$.

By our assumption on f, $\lim_{a \to 0} D_a g = D_0 g = 0$, so there exists a $\delta \in]0, \infty[$ such that for all $a \in B_A(0, 2\,\delta) \subseteq U_0$ we have $\|D_a g\|_\infty \leq \frac{1}{2}$. Hence by Lemma (5.5.5), for any $a \in \overline{B_A(0, \delta)}$ we have $\|g(a)\|_A = \|g(a) - g(0)\|_A \leq \frac{1}{2} \|a\|_A \leq \frac{\delta}{2}$, so $g(\overline{B_A(0, \delta)}) \subseteq \overline{B_A(0, \delta/2)}$.

Now let $b \in \overline{B_A(0, \delta/2)}$ and define

$$g_b : U_0 \to A : a \mapsto g(a) + b = a - f(a) + b$$

then $g_b(a) = a$ if and only if $f(a) = b$.

Let $a \in \overline{B_A(0, \delta)}$, then $\|g_b(a)\|_A = \|g(a) + b\|_A \leq \|g(a)\|_A + \|b\|_A \leq \delta/2 + \delta/2 = \delta$. Hence $g_b(\overline{B_A(0, \delta)}) \subseteq \overline{B_A(0, \delta)}$. As $D_a g_b = D_a g$, $\|D_a g_b\| \leq \frac{1}{2}$ for all $a \in B_A(0, 2\delta)$, so by Lemma (5.5.5), g_b is $\frac{1}{2}$-**c**. Now $\overline{B_A(0, \delta)} \subseteq A$ is closed and A is complete, so by Lemma (2.5.18) $\overline{B_A(0, \delta)}$ is complete and by Theorem (2.5.21), there exists a unique $a \in \overline{B_A(0, \delta)}$ such that $g_b(a) = a$, that is, such that $f(a) = b$.

So for all $b \in \overline{B_A(0, \delta/2)}$ there exists a unique $a \in \overline{B_A(0, \delta)}$ such that $f(a) = b$. Let $V := B_A(0, \delta/2)$ and define $h : V \to \overline{B(0, \delta)}$ by $h(b) := a$ whenever $f(a) = b$ (by the preceding we know that this makes h well-defined). Now for any $a_1, a_2 \in \overline{B_A(0, \delta)}$ we have that $\|a_1 - a_2\|_A = \|f(a_1) + g(a_1) - f(a_2) - g(a_2)\|_A \leq \|f(a_1) - f(a_2)\|_A + \|g(a_1) - g(a_2)\|_A \leq \|f(a_1) - f(a_2)\| + \frac{1}{2} \|a_1 - a_2\|_A$, so $\|a_1 - a_2\|_A \leq 2 \|f(a_1) - f(a_2)\|_A$ and hence

$$\|h(b_1) - h(b_2)\|_A \leq 2 \|b_1 - b_2\|_A$$

for all $b_1, b_2 \in V$. This makes h **c**.

By letting $U := h(V) = f^{-1}(V) \subseteq A$ which is an open neighbourhood of 0 as f **c** and $f(0) = 0$, we see that $h : V \to U$ **c** satisfies $h = (f|_U)^{-1}$.

Now we need to show that $h \in C^k(V, A)$. Since $D_0 f \in L(A, A)^*$ (which is open in $L(A, A)$ by Theorem (5.5.6)) there exists an $\epsilon \in]0, \infty[$ such that $B_{L(A,A)}(D_0 f, 2\,\epsilon) \subseteq L(A, A)^*$. Furthermore, as $a \mapsto D_a f$ **c**, we can choose the δ we established earlier smaller such that also $D_a f \in B_{L(A,A)}(D_0 f, \epsilon) \subseteq L(A, A)^*$ for all $a \in B_A(0, 2\,\delta)$. By these choices, $D_a f : A \to A$ is bijective for all $a \in B_A(0, 2\,\delta)$. Furthermore, since $f \in C^k(U_0, B)$, the map $B_A(0, 2\,\delta) \times A \to A : (a, a_1) \mapsto D_a f(a_1)$ is C^{k-1}. The map $B_A(0, 2\,\delta) \to L(A, A)^* : a \mapsto D_a f$ **c** by assumption and since inversion $L(A, A)^* \to L(A, A)^* : h \mapsto h^{-1}$ **c** by Theorem

(5.5.6), we have that $B_A(0, 2\,\delta) \to L(A, A)^* : a \mapsto [D_a f]^{-1}$ $\textcircled{c}$. By Theorem (5.5.6) the map $B_A(0, 2\,\delta) \times A \to A : (a, a_1) \mapsto [D_a f]^{-1}(a_1)$ is therefore $\textcircled{c}$. Now with Theorem (5.2.6) we find that $B_A(0, 2\,\delta) \times A \to A : (a, a_1) \mapsto [D_a f]^{-1}(a_1)$ is C^{k-1}, because $(a, a_1) \mapsto D_a f(a_1)$ is C^{k-1}.

Consider the map $i : B_A(0, 2\,\delta) \times B_A(0, 2\,\delta) \to L(A, A)$ given by

$$i(a_1, a_2) := \Big(A \to A : a_3 \mapsto \int_0^1 \Big(\alpha \mapsto D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_3)\Big)\Big).$$

Then with Theorem (5.3.5) and the fact that for all $\alpha \in [0, 1]$ we have $\|\alpha\, a_2 + (1-\alpha)\, a_1\|_A \leq \alpha\, \|a_2\| + (1-\alpha)\, \|a_1\|_A < (\alpha + 1 - \alpha)\, 2\,\delta = 2\,\delta$, we find

$$\begin{aligned}
\|(D_0 f - i(a_1, a_2))(a_3)\|_A &= \left\| D_0 f(a_3) - \int_0^1 \Big(\alpha \mapsto D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_3)\Big) \right\|_A \\
&= \left\| \int_0^1 \Big(\alpha \mapsto D_0 f(a_3) - D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_3)\Big) \right\|_A \\
&\leq \int_0^1 \Big(\alpha \mapsto \|D_0 f(a_3) - D_{\alpha\, a_2 + (1-\alpha)\, a_1} f(a_3)\|_A\Big) \\
&\leq \int_0^1 \Big(\alpha \mapsto \|D_0 f - D_{\alpha\, a_2 + (1-\alpha)\, a_1} f\|_\infty\, \|a_3\|_A\Big) \\
&\leq \int_0^1 \Big(\alpha \mapsto \epsilon\, \|a_3\|_A\Big) \\
&= \epsilon\, \|a_3\|_A.
\end{aligned}$$

Hence $\|i(a_1, a_2)\|_\infty \leq \epsilon < 2\,\epsilon$, so $i(a_1, a_2) \in B_{L(A,A)}(D_0 f, 2\,\epsilon) \subseteq L(A, A)^*$ for all $a_1, a_2 \in B_A(0, 2\,\delta)$. As $B_A(0, 2\,\delta) \to L(A, A)^* : a \mapsto D_a f$ $\textcircled{c}$, we find with Lemma (5.3.7) that i $\textcircled{c}$, which in turn by Theorem (5.5.6) gives us that $(a_1, a_2, a_3) \mapsto i(a_1, a_2)(a_3)$ $\textcircled{c}$.

Let $a_1, a_2 \in U$ and $b_1 := f(a_1) \in V$, $b_2 := f(a_2) \in V$, then (use Theorem (5.3.13), $f \in C^1(U, A)$)

$$\begin{aligned}
b_2 - b_1 &= f(a_2) - f(a_1) \\
&= i(a_1, a_2)(a_2 - a_1) \\
&= i(h(b_1), h(b_2))(h(b_2) - h(b_1)).
\end{aligned}$$

Now as $a_1, a_2 \in U \subseteq B_A(0, 2\,\delta)$, we have that $i(a_1, a_2) \in L(A, A)^*$ is invertible. Let us therefore define the map $j : B_A(0, 2\,\delta) \times B_A(0, 2\,\delta) \to L(A, A)$ by

$$j(a_1, a_2) := [i(a_1, a_2)]^{-1}$$

then j $\textcircled{c}$ since inversion is continuous by Theorem (5.5.6). Applying j on both sides we find for all $b_1, b_2 \in V$ that

$$h(b_2) - h(b_1) = j(h(b_1), h(b_2))(b_2 - b_1).$$

As h, j $\textcircled{c}$, $(b_1, b_2, b_3) \mapsto j(h(b_1), h(b_2))(b_3)$ $\textcircled{c}$ by Theorem (5.5.6), furthermore, this map is linear in b_3. Hence by Theorem (5.3.13), $h \in C^1(V, A)$ and

$$D_b h(b_1) = j(h(b), h(b))(b_1) = [i(h(b), h(b))]^{-1}(b_1) = [D_{h(b)} f]^{-1}(b_1).$$

We already established that $U \times A \to A : (a, a_1) \mapsto [D_a f]^{-1}(a_1)$ is C^{k-1}, so using induction, Equation (5.4), and the fact that $h \in C^1(V, A)$ with derivative $D_b h(b_1) = [D_{h(b)} f]^{-1}(b_1)$, we find that $V \times A \to A : (b, b_1) \mapsto [D_{h(b)} f]^{-1}(b_1)$ is C^{k-1} and therefore that $h \in C^k(V, A)$.

Therefore $f|_U : U \to V$ is a C^k diffeomorphism with inverse h. □

Theorem 5.5.9: Implicit function theorem

Let A, B, C /$\mathbb{K}$, $U_0 \subseteq A$, $V_0 \subseteq B$ open. Let $f \in C^k(U_0 \times V_0, C)$ such that $U_0 \times V_0 \to L(A \times B, C) : (a, b) \mapsto D_{(a,b)} f$ (automatically true for $k \geq 2$ by Corollary (5.5.7)).

Define for $a \in U$ and $b \in V$ the maps

$$f_a : V_0 \to C : b_1 \mapsto f(a, b_1), \qquad f_b : U_0 \to C : a_1 \mapsto f(a_1, b).$$

Suppose that for a certain $(a_0, b_0) \in U_0 \times V_0$ and $c_0 \in C$ we have $f(a_0, b_0) = c_0$ and that $D_{a_0} f_{b_0} : A \to C$ is bijective.

Then there exists an open neighbourhood $U \subseteq U_0$ of a_0 in A and $V \subseteq V_0$ of b_0 in B and a map $h : V \to U$ such that

- $h \in C^k(V, A)$,
- for all $b \in V$, $h(b) \in U$ is the unique element in U for which $f(h(b), b) = c_0$,
- for all $b \in V$,

$$D_b h(b_1) = -[D_{h(b)} f_b]^{-1}(D_b f_{h(b)}(b_1)). \tag{5.9}$$

Proof. We follow [DK2004I]. By Corollary (4.4.6) $D_{a_0} f_{b_0} : A \to C$ is a -isomorphism. Hence we can consider the map

$$(U_0 - a_0) \times (V_0 - b_0) \to [D_{a_0} f_{b_0}]^{-1}(f(U_0, V_0)) : (a, b) \mapsto [D_{a_0} f_{b_0}]^{-1}(f(a_0 + a, b_0 + b) - c_0)$$

which is C^k, instead of f. Therefore we may suppose that $C = A$, $U_0 \ni a_0 = 0$, $V_0 \ni b_0 = 0$, $c_0 = 0$, and $D_{a_0} f_{b_0} = \mathrm{id}_A$.

Now define the map $g : U_0 \times V_0 \to A \times B$ by $g(a, b) := (f(a, b), b)$. Then by Lemma (5.1.11) and Lemma (5.1.13) we have that $g \in C^k(U_0 \times V_0, A \times B)$ with $D_{(a,b)} g(a_1, b_1) = (D_{(a,b)} f(a_1, b_1), b_1) = (D_a f_b(a_1) + D_b f_a(b_1), b_1)$. In particular $D_{(a_0,b_0)} g(a_1, b_1) = (a_1 + D_{b_0} f_{a_0}(b_1), b_1)$, so $D_{(a_0,b_0)} g : A \times B \to A \times B$ is bijective. Also $(a, b) \mapsto D_{(a,b)} g$ as $(a, b) \mapsto D_{(a,b)} f$, therefore by Theorem (5.5.8) there exist open neighbourhoods $U, U' \subseteq U_0$ and $V, V' \subseteq V_0$ of 0 in A and B respectively, and a C^k diffeomorphism $i : U' \times V \to U \times V'$ which is an inverse of $g|_{U \times V'}$.

Since $g(a, b) = (f(a, b), b)$ we can write $i(c, b) = (h(c, b), b)$ for $h : U' \times V \to U$, C^k, and take $V' = V$. Now for $a \in U$, $b \in V$, $c \in C$ we have $f(a, b) = c$ if and only if $g(a, b) = (c, b)$ if and only if $(a, b) = i(c, b)$ if and only if $h(c, b) = a$. Hence $f(a, b) = 0$ if and only if $h(0, b) = a$, for all $(a, b) \in U \times V$. Therefore $V \to U : b \mapsto h(0, b)$ is the C^k map we seek, the derivative of which follows from $g \circ i = \mathrm{id}_{U' \times V}$: $(c, b) = g(i(c, b)) = g(h(c, b), b) = (f(h(c, b), b), b)$, so using Theorem (5.1.8) and Lemma (5.1.13) we find that $c_1 = D_{(h(c,b),b)} f(D_{(c,b)} h(c_1, b_1), b_1) = D_{h(c,b)} f_b(D_{(c,b)} h(c_1, b_1)) + D_b f_{h(c,b)}(b_1)$. So in particular for $c = c_1 = 0$ we have $D_{h(0,b)} f_b(D_{(0,b)} h(0, b_1)) = -D_b f_{h(c,b)}(b_1)$ which yields Equation (5.9). □

Theorem 5.5.10: Existence and uniqueness of solutions of ordinary differential equations ($e^{\cdot\, f}$)
Let A /$\mathbb{K}$, $U \subseteq A$ open. Let $f \in C^k(U, A)$ such that $U \to L(A, A) : a \mapsto D_a f$ (automatically true for $k \geq 2$ by Corollary (5.5.7)).

Then there exists an open set $V \subseteq \mathbb{R} \times A$ for which

$$\{0\} \times U \subseteq V \subseteq \mathbb{R} \times U,$$

and there exists a map, called the *flow of* f,

$$e^{\cdot\, f} : V \to U : (\alpha, a) \mapsto e^{\alpha\, f}(a)$$

which satisfies $e^{\cdot\, f} \in C^k(V, A)$.

For convenience we define for all $a \in U$,

$$S(a) := \{\alpha \in \mathbb{R} \mid (\alpha, a) \in V\} \subseteq \mathbb{R},$$

and for all $\alpha \in \mathbb{R}$,

$$U(\alpha) := \{a \in U \mid (\alpha, a) \in V\} \subseteq U.$$

The flow $e^{\cdot\, f}$ has the following properties.

- For all $\alpha \in \mathbb{R}$, the map

  $$U(\alpha) \to U : a \mapsto e^{\alpha\, f}(a)$$

  is C^k and satisfies for all $\alpha, \beta \in \mathbb{R}$

  $$e^{0\, f} = \mathrm{id}_U, \qquad e^{(\alpha+\beta)\, f} = e^{\alpha\, f} \circ e^{\beta\, f}$$

  whenever the right-hand-side is well-defined.

- For all $a \in U$, $0 \in S(a) \subseteq \mathbb{R}$ is an open interval and the map

  $$g_a : S(a) \to U : \alpha \mapsto e^{\alpha\, f}(a)$$

  satisfies

  $$g_a \in C^k(S(a), A), \qquad g_a(0) = a, \qquad \forall \alpha \in S(a) : g_a'(\alpha) = f(g_a(\alpha)). \tag{5.10}$$

  This g_a is furthermore the unique and maximal solution to Equation (5.10) in the sense that if there is another map $h : S \to U$, with $S \subseteq \mathbb{R}$ an open interval, satisfying $h'(\alpha) = f(h(\alpha))$ for all $\alpha \in S$ and $h(\alpha_0) = a$, then $S \subseteq (\alpha_0 + S(a))$ and $h(\alpha) = g_a(\alpha - \alpha_0)$ for all $\alpha \in S$.

Proof. We follow [DK2000] and make extensive use of Corollary (5.1.7).

First we show existence of solutions to the equations $g(0) = a$, $g'(\alpha) = f(g(\alpha))$. Fix $a_0 \in U$. Since f and $a \mapsto D_a f$ there exists an $\epsilon \in]0, \infty[$ such that $B_A(a_0, 2\,\epsilon) \subseteq U$, and for all $a \in B_A(a_0, 2\,\epsilon)$ we have $\|f(a) - f(a_0)\|_A < 1$ and $\|D_a f - D_{a_0} f\|_\infty < 1$, hence

$$\|f(a)\|_A \leq \|f(a_0)\|_A + 1 =: \alpha_0, \qquad \|D_a f\|_\infty \leq \|D_{a_0} f\|_\infty + 1 =: \alpha_1$$

for all $a \in B_A(a_0, 2\,\epsilon)$. Pick $\delta \in]0, \infty[$ such that $\alpha_0\,\delta < \frac{\epsilon}{2}$, and $\alpha_1\,\delta < 1$. Let

$$B := \{g :]-\delta, \delta[\to \overline{B_A(a_0, \epsilon)} \subseteq U \mid \|g\|_B < \infty\}$$

together with the norm

$$\|g\|_B := \sup\{\|g(\alpha)\|_A \in \mathbb{R} \mid \alpha \in]-\delta, \delta[\}.$$

Then B Ba $/\mathbb{K}$ ([Mun2000], Theorem 43.6).

Let $F : B_A(a_0, \frac{\epsilon}{2}) \times B \to B$ be defined by

$$F(a, g)(\alpha) := a + \int_0^\alpha \Big(\beta \mapsto f(g(\beta))\Big).$$

Note that for $g \in B$, $F(a, g) = g$ if and only if $g(0) = a$ and for all α, $g'(\alpha) = f(g(\alpha))$ by Theorem (5.3.8).

Then for any $a, a' \in B_A(a_0, \frac{\epsilon}{2})$, $\alpha \in]-\delta, \delta[$ we have (as $g(\alpha) \in \overline{B_A(a_0, \epsilon)} \subseteq B_A(a_0, 2\,\epsilon)$)

$$\begin{aligned}
\|F(a, g)(\alpha) - a'\|_A &= \left\| a + \int_0^\alpha \Big(\beta \mapsto f(g(\beta))\Big) - a' \right\|_A \\
&\leq \|a - a'\|_A + \left| \int_0^\alpha \Big(\beta \mapsto \|f(g(\beta))\|_A\Big) \right| \\
&\leq \|a - a'\|_A + |\alpha|\,\alpha_0 \\
&< \|a - a'\|_A + \alpha_0\,\delta,
\end{aligned}$$

so in particular by our choice of δ, $F(a, g)(\alpha) \in B_A(a_0, \epsilon) \subseteq U$ for all $a \in B_A(a_0, \frac{\epsilon}{2})$, $\alpha \in]-\delta, \delta[$, which makes $F(a, g) \in B$. By Lemma (5.5.5), $\|f(a') - f(a)\|_A \leq \alpha_1\,\|a' - a\|_A$, so for $g, h \in B$

$$\begin{aligned}
\|F(a, g)(\alpha) - F(a', h)(\alpha)\|_A &= \left\| a + \int_0^\alpha \Big(\beta \mapsto f(g(\beta))\Big) - a' - \int_0^\alpha \Big(\beta \mapsto f(h(\beta))\Big) \right\|_A \\
&\leq \|a - a'\|_A + \left| \int_0^\alpha \Big(\beta \mapsto \|f(g(\beta)) - f(h(\beta))\|_A\Big) \right| \\
&\leq \|a - a'\|_A + \left| \int_0^\alpha \Big(\beta \mapsto \alpha_1\,\|g(\beta) - h(\beta)\|_A\Big) \right| \\
&\leq \|a - a'\|_A + |\alpha|\,\alpha_1\,\|g - h\|_B \\
&< \|a - a'\|_A + \alpha_1\,\delta\,\|g - h\|_B.
\end{aligned}$$

Hence

$$\begin{aligned}
\|F(a, g) - a'\|_B &< \|a - a'\|_A + \alpha_0\,\delta, \\
\|F(a, g) - F(a', h)\|_B &< \|a - a'\|_A + \alpha_1\,\delta\,\|g - h\|_B
\end{aligned} \tag{5.11}$$

for all $a, a' \in B_A(a_0, \frac{\epsilon}{2})$ and $g, h \in B$.

Let $a \in B_A(a_0, \frac{\epsilon}{2})$, then for all $g \in \overline{B_B(a_0, \epsilon)}$ we have $\|F(a, g) - a_0\|_B \leq \frac{\epsilon}{2} + \alpha_0\,\delta < \epsilon$, so $F(a, g) \in \overline{B_B(a_0, \epsilon)}$. Also, for any $g, h \in \overline{B_B(a_0, \epsilon)}$ we have $\|F(a, g) - F(a, h)\|_B \leq \alpha_1\,\delta\,\|g - h\|_B$ where $\alpha_1\,\delta < 1$. Now as $\overline{B_B(a_0, \epsilon)} \subseteq B$ is closed, it is complete by Lemma (2.5.18) and therefore by Theorem (2.5.21), there exists a unique $g_a \in \overline{B_B(a_0, \epsilon)}$ such that $F(a, g_a) = g_a$. By Theorem (5.3.8) we have that $g_a \in C^1(]-\delta, \delta[, A)$, $g_a(0) = a + 0$, and $g_a'(\alpha) = 0 + f(g_a(\alpha))$. This function g_a is by Theorem (5.3.8) the solution $g_a :]-\delta, \delta[\to A$ in B to $g_a(0) = a$ and $g_a'(\alpha) = f(g_a(\alpha))$ for all α, and furthermore the unique solution in $\overline{B_B(a_0, \epsilon)}$ by Theorem (2.5.21).

From the definition of F, $f \in C^k(U, A)$, and Lemma (5.3.7) we find that $F \in C^k(B_A(a_0, \frac{\epsilon}{2}) \times B, B)$ and furthermore for all $a \in B_A(a_0, \frac{\epsilon}{2})$ that by Equation (5.11),

$$\left\| D_g\Big(h \mapsto F(a, h)\Big) \right\|_\infty \leq 0 + \alpha_1\, \delta < 1.$$

Hence by Theorem (5.5.6) the map $(\mathrm{id}_B - D_g(h \mapsto F(a, h))) : B \to B$ is invertible. This means that the map $B \times B_A(a_0, \frac{\epsilon}{2}) \to B : (g, a) \mapsto g - F(a, g)$ satisfies the conditions of Theorem (5.5.9) at $(g_a, a) \in B \times U$ for all $a \in B_A(a_0, \frac{\epsilon}{2})$. Hence g_a depends in a C^k fashion on a, that is, for all $a \in B_A(a_0, \frac{\epsilon}{2})$ there exists a $\delta' \in]0, \frac{\epsilon}{2}[$ such that the map $B_A(a, \delta') \to B : a' \mapsto g_{a'}$ is C^k.

So for all $a_0 \in U$ there exist $\delta, \delta', \epsilon \in]0, \infty[$ such that for all $a \in B_A(a_0, \frac{\epsilon}{2})$ there exists a unique $g_a \in \overline{B_B(a_0, \epsilon)}$, $g_a \in C^1(] - \delta, \delta[, A)$ with $g_a(0) = a$ and $g_a'(\alpha) = f(g(\alpha))$ for all $\alpha \in] - \delta, \delta[$. Furthermore, the map $B_A(a_0, \delta') \to B : a \mapsto g_a$ is C^k.

Now we will show that such solutions g_a are unique in a global sense. Suppose we have two maps $g : S \to U$ C^1 and $h : T \to U$ C^1, where $S, T \subseteq \mathbb{R}$ are open intervals. Suppose g and h satisfy $g'(\alpha) = f(g(\alpha))$ and $h'(\beta) = f(h(\beta))$ for all $\alpha \in S$, $\beta \in T$ and $g(\alpha_0) = h(\beta_0)$ for some $\alpha_0 \in S$, $\beta_0 \in T$. Then first of all, by Theorem (5.1.8), the maps $i : (S - \alpha_0) \to U : \alpha \mapsto g(\alpha - \alpha_0)$, $j : (T - \beta_0) \to U : \beta \mapsto h(\beta - \beta_0)$ satisfy $i(0) = j(0)$ and $i'(\alpha) = f(i(\alpha))$, $j'(\beta) = f(j(\beta))$. Let $a_0 := i(0) = j(0) \in U$, then by the preceding there exist $\delta, \delta', \epsilon \in]0, \infty[$ and a unique $g_{a_0} :] - \delta, \delta[\to U$ in $\overline{B_B(a_0, \epsilon)}$ with $g_{a_0}(0) = a_0$, $g_{a_0}'(\alpha) = f(g_{a_0}(\alpha))$. We can furthermore choose δ smaller such that $] - \delta, \delta[\subseteq [-\delta, \delta] \subseteq (S - \alpha_0) \cap (T - \beta_0)$ since $0 \in (S - \alpha_0) \cap (T - \beta_0)$ is an open interval. As i, j, $\lim_{\alpha \to 0} i(\alpha) = \lim_{\alpha \to 0} j(\alpha) = a_0$, and $[-\delta, \delta]$, so $\|i([-\delta, \delta])\|_A$ and $\|j([-\delta, \delta])\|_A$ are bounded in $\mathbb{R}$. Hence we can take δ smaller, such that $i|_{]-\delta,\delta[}, j|_{]-\delta,\delta[} \in \overline{B_B(a_0, \epsilon)}$. Then, as g_{a_0} is the unique solution to $F(a_0, g_{a_0}) = g_{a_0}$ and $i|_{]-\delta,\delta[} = F(a_0, i|_{]-\delta,\delta[})$, we find that $g_{a_0} = i|_{]-\delta,\delta[}$ and similarly $g_{a_0} = j|_{]-\delta,\delta[}$.

Hence there exists a $\delta \in]0, \infty[$ such that $i(\alpha) = j(\alpha)$ for all $\alpha \in] - \delta, \delta[$.

Now consider the collection $S' := \{\alpha \in (S - \alpha_0) \cap (T - \beta_0) \mid i(\alpha) = j(\alpha)\} \subseteq (S - \alpha_0) \cap (T - \alpha_0)$. Then by the previous we know that $0 \in S'$ implies that there exists a $\delta \in]0, \infty[$ such that $] - \delta, \delta[\subseteq S'$. By applying the same argument again, we find that for any $\alpha \in S'$ there exists a $\delta_\alpha \in]0, \infty[$ such that $]\alpha - \delta_\alpha, \alpha + \delta_\alpha[\subseteq S'$. Hence $S' \subseteq (S - \alpha_0) \cap (T - \beta_0)$ is open. On the other hand, S' is the inverse image of $\{0\}$ of the map $((a, a') \mapsto a - a') \circ (\alpha \mapsto (i(\alpha), j(\alpha)))$ which is and therefore (Lemma (2.1.14)) S' is closed. So S' is a subset of an open interval that is both open and closed, hence S' is either empty or equal to the entire open interval. Because $0 \in S'$ we find that necessarily $S' = (S - \alpha_0) \cap (T - \beta_0)$. Therefore $g(\alpha_0 + \alpha) = h(\beta_0 + \alpha)$ for all $\alpha \in (S - \alpha_0) \cap (T - \beta_0)$.

So for any two paths $g : S \to U$, $h : T \to U$ both C^1 and satisfying $g'(\alpha) = f(g(\alpha))$, $h'(\beta) = f(h(\beta))$ for all $\alpha \in S$, $\beta \in T$ we have that if $g(\alpha_0) = h(\beta_0)$ for some $\alpha_0 \in S$, $\beta_0 \in T$, then $g(\alpha_0 + \alpha) = h(\beta_0 + \alpha)$ for all $\alpha \in (S - \alpha_0) \cap (T - \beta_0)$.

This uniqueness property can be used to increase the domain $] - \delta, \delta[$ of our solutions g_a. Let for $a_0 \in U$, $S(a_0)$ denote the union of all open intervals $S \subseteq \mathbb{R}$

containing 0 for which there exists a $g : S \to U$ that is C^1 with $g(0) = a_0$ and $g'(\alpha) = f(g(\alpha))$ for all $\alpha \in S$. In particular there exists a $\delta \in]0, \infty[$ such that $] - \delta, \delta[\subseteq S(a_0)$ by construction of g_{a_0}, so $S(a_0) \subseteq \mathbb{R}$ is an open interval containing 0.

With a slight abuse of notation, we will write g_{a_0} for the maximal C^1 curve defined on $S(a_0)$ as the union of all curves whose domain is contained in the union $S(a_0)$. This makes g_{a_0} well-defined because of the uniqueness property stated above.

It is clear that with this definition, $g_{a_0} : S(a_0) \to U$ is the unique and maximal solution from the second point of the theorem.

Let $\alpha \in S(a)$ and $\beta \in S(g_a(\alpha))$. Then as $g_a(\alpha) = g_{g_a(\alpha)}(0)$ and $\beta \mapsto g_a(\alpha+\beta)$, $g_{g_a(\alpha)}$ are both maximal solutions to Equation (5.10), we have $S(\alpha) = \alpha + S(g_a(\alpha))$ and $g_{g_a(\alpha)}(\beta) = g_a(\alpha + \beta)$.

In particular $\alpha + \beta \in S(a)$ for all $\alpha \in S(a)$ and $\beta \in S(g_a(\alpha))$.

We are now going to construct V and $e^{\cdot f}$.

Choose

$$V := \bigcup \{S(a_0) \times \{a_0\} \mid a_0 \in U\} \subseteq \mathbb{R} \times U.$$

Let $a_0 \in U$, $\alpha \in S(a_0)$ and choose $a_1 := g_{a_0}(\alpha) \in U$.

By our first assertion there exist $\epsilon, \delta, \delta' \in]0, \infty[$ such that for all $a \in B_A(a_0, \frac{\epsilon}{2})$ and all $a \in B_A(a_1, \frac{\epsilon}{2})$ we have $] - \delta, \delta[\subseteq S(a)$ and that $B_A(a_0, \delta') \to B : a \mapsto g_a$ is C^k. Choose $\delta' \leq \delta \leq \frac{\epsilon}{2}$ for convenience.

We have that $\lim_{a \to a_0} g_a = g_{a_0}$ in B, so there exists a $\delta'' \in]0, \delta']$ such that for all $a \in B_A(a_0, \delta'')$ we have $\|g_a - g_{a_0}\|_B < \delta'$. In particular for $a \in B_A(a_0, \delta'')$ we have $\|g_a(\alpha) - a_1\|_A \leq \|g_a - g_{a_0}\|_B < \delta'$, so $g_a(\alpha) \in B_A(a_1, \delta') \subseteq B_A(a_1, \frac{\epsilon}{2})$ and hence $] - \delta, \delta[\subseteq S(g_a(\alpha))$. So for $a \in B_A(a_0, \delta'')$, $\alpha \in S(a)$ and $] - \delta, \delta[\subseteq S(g_a(\alpha))$, hence $]\alpha - \delta, \alpha + \delta[\subseteq S(a)$. But this means that $(\alpha, a_0) \in]\alpha - \delta, \alpha + \delta[\times B_A(a_0, \delta'') \subseteq V$: V is open.

Since $\alpha = 0 \in S(a_0)$ for all $a_0 \in U$, we furthermore find that $\{0\} \times U \subseteq V$.

This permits us to define

$$e^{\cdot f} : V \to U : (\alpha, a) \mapsto e^{\alpha f}(a) := g_a(\alpha)$$

in accordance with the notation for g_a from the theorem, which is well-defined because of uniqueness property of the g_a.

Note that $e^{0 f}(a) = g_a(0) = a = \mathrm{id}_U(a)$, so $e^{0 f} = \mathrm{id}_U$. Furthermore, whenever $(\beta, a), (\alpha, e^{\beta f}(a)) \in V$, we have $e^{\alpha f}(e^{\beta f}(a)) = e^{\alpha f}(g_a(\beta)) = g_{g_a(\beta)}(\alpha) = g_a(\alpha + \beta) = e^{(\alpha+\beta) f}(a)$.

$e^{\cdot f}$ is C^k by Theorem (5.1.8), the composition rule $e^{\alpha f} \circ e^{\beta f} = e^{(\alpha+\beta) f}$, and the fact that $a \mapsto g_a$ is C^k. □

Example 5.5.11: Notation of Theorem (5.5.10).

The notation $e^{\cdot f}$ in Theorem (5.5.10) has purposefully been introduced because of the following. Let A Ba $/\mathbb{K}$ and consider id_A which is C^∞.

Hence $e^{\cdot\, \mathrm{id}_A} : V \to A$ exists and we can study its form by looking at the map F from the proof. Fix any $a \in A$, then

$$F(a, 0)(\alpha) = a + \int_0^\alpha \Big(\beta \mapsto \mathrm{id}_A(0)\Big) = a$$

iterating F to find the solution (as is done in Theorem (2.5.21) which is used to find the solutions in the proof) we find

$$F(a, F(a,0))(\alpha) = a + \int_0^\alpha \Big(\beta \mapsto \mathrm{id}_A(a)\Big) = a + \alpha\, a.$$

So after k iterations we find

$$\underbrace{F(a, \ldots, F(a,0)\ldots)}_{k}(\alpha) = a + \alpha\, a + \frac{\alpha^2}{2}\, a + \ldots + \frac{\alpha^k}{k!}\, a$$

which is exactly the k-th order Taylor expansion of the map $\alpha \mapsto e^\alpha\, a$. Indeed this map, defined for all $\alpha \in \mathbb{R}$, is the sought-after solution since $e^0\, a = a$ and $\frac{\mathrm{d}}{\mathrm{d}\alpha} e^\alpha\, a = e^\alpha\, a = \mathrm{id}_A(e^\alpha\, a)$. Hence for all $a \in A$ we have

$$e^{\alpha\, \mathrm{id}_A}(a) = e^\alpha\, a$$

which is indeed a C^∞ map, which is furthermore defined on the entire $V = \mathbb{R} \times A$. This motivates the notation of Theorem (5.5.10).

The following lemma may be used to calculate the derivatives of the flow $e^{\cdot f}$ of a given function f.

Lemma 5.5.12: Derivatives of $e^{\cdot f}$
Let A Ba $/\mathbb{K}$, $U \subseteq A$ open, $k \in \mathbb{N}$, and $f \in C^k(U, A)$ satisfying the conditions of Theorem (5.5.10).

Let $e^{\cdot f} : V \to U$, with $V \subseteq \mathbb{R} \times U$ open, denote the flow of f.

Then for all $0 \leq l < k$, $a \in U$, and $u_1, \ldots, u_l \in A$, the curve $g : S(a) \to A$ defined by

$$g(\alpha) := D_a^l e^{\alpha\, f}(u_1, \ldots, u_l)$$

satisfies for all $\alpha \in S(a)$

$$g'(\alpha) = D_a^l (f \circ e^{\alpha\, f})(u_1, \ldots, u_l)$$
$$g(0) = \begin{cases} a & l = 0 \\ u_1 & l = 1 \\ 0 & l > 1. \end{cases}$$

In particular, the flow of

$$U \to A : a \mapsto D_a^l (f \circ e^{\alpha\, f})(u_1, \ldots, u_l)$$

gives us information about the l-th derivative of the flow of f in the directions $u_1, \ldots, u_l$.

Proof. We will use Theorem (5.5.10) extensively in this proof, particularly the fact that $e^{\cdot f}$ is C^k and that therefore all limits of quotients involving $e^{\cdot f}$ exist. Suppose $l = 0$, then $g = g_a$, so $g(0) = g_a(0) = a$ and $g'(\alpha) = g_a'(\alpha) = f(g_a(\alpha)) = f(e^{\alpha\, f}(a)) = D_a^0 (f \circ e^{\alpha\, f})$ by Equation (5.10).

Suppose $l = 1$, then $g(0) = D_a e^{0\,f}(u_1) = D_a \operatorname{id}_U(u_1) = \operatorname{id}_U(u_1) = u_1$. Furthermore with Corollary (5.1.7) and Theorem (5.1.16)

$$\begin{aligned}
g'(\alpha) &= \lim_{\beta\to 0}\frac{1}{\beta}\Big(D_a e^{(\alpha+\beta)\,f}(u_1) - D_a e^{\alpha\,f}(u_1)\Big)\\
&= D^2_{(\alpha,a)} e^{\cdot\,f}((0,u_1),(1,0))\\
&= D^2_{(\alpha,a)} e^{\cdot\,f}((1,0),(0,u_1))\\
&= D_{(\alpha,a)}\Big((\alpha',a') \mapsto D_{(\alpha',a')}\Big((\alpha'',a'') \mapsto e^{\alpha''\,f}(a'')\Big)(1,0)\Big)(0,u_1)\\
&= D_{(\alpha,a)}\Big((\alpha',a') \mapsto \lim_{\beta\to 0}\frac{1}{\beta}\Big(g_{a'}(\alpha'+\beta) - g_{a'}(\alpha')\Big)\Big)(0,u_1)\\
&= D_{(\alpha,a)}\Big((\alpha',a') \mapsto g'_{a'}(\alpha')\Big)(0,u_1)\\
&\overset{(5.10)}{=} D_{(\alpha,a)}\Big((\alpha',a') \mapsto f(g_{a'}(\alpha'))\Big)(0,u_1)\\
&= D_{(\alpha,a)}\Big((\alpha',a') \mapsto f(e^{\alpha'\,f}(a'))\Big)(0,u_1)\\
&= D_a(f \circ e^{\alpha\,f})(u_1).
\end{aligned}$$

Suppose $l > 1$, then $g(0) = D^l_a e^{0\,f}(u_1,\dots,u_l) = D^l_a \operatorname{id}_U(u_1,\dots,u_l) = 0$. Furthermore by the above calculation, we may use induction to assume that for $l-1$

$$D^l_{(\alpha,a)} e^{\cdot\,f}((0,u_1),\dots,(0,u_{l-1}),(1,0)) = D^{l-1}_a(f \circ e^{\alpha\,f})(u_1,\dots,u_{l-1}).$$

Hence

$$\begin{aligned}
D^l_a(f \circ e^{\alpha\,f})(u_1,\dots,u_l) &= D^{l+1}_{(\alpha,a)} e^{\cdot\,f}((0,u_1),\dots,(0,u_{l-1}),(1,0),(0,u_l))\\
&= D^{l+1}_{(\alpha,a)} e^{\cdot\,f}((0,u_1),\dots,(0,u_{l-1}),(0,u_l),(1,0))\\
&= \lim_{\beta\to 0}\frac{1}{\beta}\Big(D^l_{(\alpha+\beta\,1,a)} e^{\cdot\,f}((0,u_1),\dots,(0,u_l))\\
&\quad - D^l_{(\alpha,a)} e^{\cdot\,f}((0,u_1),\dots,(0,u_l))\Big)\\
&= \lim_{\beta\to 0}\frac{1}{\beta}\Big(D^l_a e^{(\alpha+\beta)\,f}(u_1,\dots,u_l) - D^l_a e^{\alpha\,f}(u_1,\dots,u_l)\Big)\\
&= g'(\alpha),
\end{aligned}$$

which shows that g has the desired property. $\square$

Chapter 6

Revisiting Christoffel's article

In this chapter we will revisit and generalise [Chr1869]. Before we start however, we will first need to introduce a few concepts that will be helpful for discussing Christoffel's article.

6.1 Preliminaries

Definition 6.1.1: k-Tensor
Let A Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, and $k \in \mathbb{N}$.

Then a *k-tensor* is a family of k-linear maps (Definition (5.2.2)) $f : U \times \underbrace{A \times \ldots \times A}_{k} \to \mathbb{K} : (a, u_1, \ldots, u_k) \mapsto f_a(u_1, \ldots, u_k)$ that is C.

We say that a k-tensor is *symmetric* if for any $\pi \in S^k$, $a \in U$, and $u_1, \ldots, u_k \in A$ we have

$$f_a(u_1, \ldots, u_k) = f_a(u_{\pi(1)}, \ldots, u_{\pi(k)}).$$

We say for $l \in \mathbb{N}$ that a k-tensor $f \in C^l(U)$ if $f \in C^l(U \times A \times \ldots \times A, \mathbb{K})$.

Definition 6.1.2: Metric
Let A Vs $/\mathbb{K}$ T2 LC, and $U \subseteq A$ open.

Then a 2-tensor $f : U \times A \times A \to \mathbb{K}$ induces a map $\hat{f} : U \times A \to A' : (a, u) \mapsto \hat{f}_a(u)$ defined by

$$\hat{f}_a : A \to A' : (u \mapsto (v \mapsto f_a(u, v))).$$

If for all $a \in U$, $\hat{f}_a$ is bijective and the map $\hat{f}^{-1} : U \times A' \to A : (a, g) \mapsto \hat{f}_a^{-1}(g)$ C, we call f *non-degenerate*.

A 2-tensor $f : U \times A \times A \to \mathbb{K}$ that is symmetric and non-degenerate is called a *metric*.

Note that any A Vs $/\mathbb{K}$ T2 LC only admits a metric if $A \simeq A'$ are Vs-isomorphic, since for any $a \in U$, $\hat{f}_a : A \to A'$ is an Vs-isomorphism if f is a metric.

Lemma 6.1.3
Let A Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, and $f : U \times A \times A \to \mathbb{K}$ a metric.

Then

- for any $g, h \in A'$ and $a \in U$ we have
$$g(\hat{f}_a^{-1}(h)) = h(\hat{f}_a^{-1}(g)),$$

- $\hat{f} : U \times A \to A'$ ⓒ, for all $a \in U$, $\hat{f}_a$ ⓛ$/\mathbb{K}$ and $\hat{f}^{-1} : U \times A' \to A$ ⓒ, $\hat{f}_a^{-1}$ ⓛ$/\mathbb{K}$.

- If $f \in C^k(U)$ for some $k \in \mathbb{N}$, then $\hat{f} \in C^k(U \times A, A')$ and $\hat{f}^{-1} \in C^k(U \times A', A)$. We then have for all $a \in U$, $u \in A$, $v, w \in A$, and $g \in A'$
$$\begin{aligned} \big(D_a \hat{f}(u)(v)\big)(w) &= D_a f(u)(v, w) \\ f_a\big(D_a \hat{f}^{-1}(u)(g), v\big) &= -D_a f(u)(\hat{f}_a^{-1}(g), v). \end{aligned} \tag{6.1}$$

Proof.

- Let $a \in U$, $g, h \in A'$. Then because f is non-degenerate:
$$g(\hat{f}_a^{-1}(h)) = (\hat{f}_a(\hat{f}_a^{-1}(g)))(\hat{f}_a^{-1}(h)) = f_a(\hat{f}_a^{-1}(g), \hat{f}_a^{-1}(h))$$
and the left hand side is symmetric if and only if the right hand side is.

- The map $\hat{f} : U \times A \to A'$ is ⓒ: $\hat{f}(a, u)(v) = \hat{f}_a(u)(v) = f_a(u, v)$ and $f : U \times A \times A \to \mathbb{K}$ ⓒ. Because for each $a \in U$, $f_a : A \times A \to \mathbb{K}$ is 2-ⓛ, $\hat{f}_a$ ⓛ$/\mathbb{K}$. $\hat{f}^{-1}$ is ⓒ by assumption and $\hat{f}_a^{-1}$ ⓛ as the inverse of a linear map.

- For $a \in U$, $u, a_1, u_1, v \in A$ and $\alpha \in \mathbb{K}$ small enough but not equal to zero, we have
$$\begin{aligned} &\frac{1}{\alpha}\Big(\hat{f}_{a+\alpha\, a_1}(u + \alpha\, u_1)(v) - \hat{f}_a(u)(v)\Big) \\ &= \frac{1}{\alpha}\Big(f_{a+\alpha\, a_1}(u + \alpha\, u_1, v) - f_a(u, v)\Big) \\ &= \frac{1}{\alpha}\Big(f_{a+\alpha\, a_1}(u, v) - f_a(u, v)\Big) + f_{a+\alpha\, a_1}(u_1, v), \end{aligned}$$
so as $f \in C^1(U)$ we see that by taking the limit for $(\alpha, (a_1, u_1)) \to (0, (a_2, u_2))$, we obtain that the above goes to
$$D_a f(a_2)(u, v) + f_a(u_2, v)$$
and since the map $(A \to \mathbb{K} : v \mapsto D_a f(a_2)(u, v) + f_a(u_2, v)) \in A'$ we therefore have by Lemma (5.1.4) that $\hat{f}$ ⓓ (a, u) with
$$D_{(a,u)} \hat{f}(a_2, u_2) = \big(v \mapsto D_a f(a_2)(u, v) + f_a(u_2, v)\big)$$
which gives us the desired formula for $D_a \hat{f}(a_2)$ if we use the notation of Definition (5.2.2). Since this expression is continuous in all its variables, $\hat{f} \in C^1(U \times A \to A')$. Now we can differentiate this expression using Theorem (5.1.8) and with induction conclude that $f \in C^k(U)$ implies that $\hat{f} \in C^k(U \times A, A')$.

  Because $\hat{f}_a : A \to A'$ is bijective for all $a \in U$, $\hat{f} \in C^k(U \times A, A')$, and the collection of inverses $\hat{f}^{-1}$ ⓒ by assumption, we can apply Theorem (5.2.6) to conclude that $\hat{f}^{-1} \in C^k(U \times A', A)$. Furthermore, Equation (5.5) gives us that for $a \in U$, $u, v \in A$, and $g \in A'$,

$$\begin{aligned} f_a(D_a\hat{f}^{-1}(u)(g), v) &= -f_a(\hat{f}_a^{-1}(D_a\hat{f}(u)(\hat{f}_a^{-1}(g))), v) \\ &= \hat{f}_a(\hat{f}_a^{-1}(D_a\hat{f}(u)(\hat{f}_a^{-1}(g))))(v) \\ &= -D_a\hat{f}(u)(\hat{f}_a^{-1}(g))(v) \\ &= -D_af(u)(\hat{f}_a^{-1}(g), v). \end{aligned}$$

□

6.2 Generalisation

We will now generalise the calculations in [Chr1869], a translation of which is included in Chapter 8. Between [Chr1869] and our calculations here, we have the following correspondences (here e_i denotes the i-th basis vector of $\mathbb{R}^n$ for $1 \leq i \leq n$).

Here	[Chr1869]
A	$A = \mathbb{R}^n$, $x' \in A$
B	$B = \mathbb{R}^n$, $x \in B$
f	$x = x(x') = f(x')$
Df	$\frac{\partial x_i}{\partial x'_j} = \frac{\partial x_i(x')}{\partial x'_j} = D_{x'}f_i(e_j)$
g	$\omega'_{ij} = \omega'_{ij}(x') = g_{x'}(e_i, e_j)$
h	$\omega_{ij} = \omega_{ij}(x) = h_x(e_i, e_j)$
B	$\begin{bmatrix} ij \\ k \end{bmatrix}(x) = B_x^h(e_i, e_j, e_k)$
Γ	$\sum_{k=1}^n \begin{Bmatrix} ij \\ k \end{Bmatrix}(x)\, e_k = \Gamma_x^h(e_i, e_j)$
R	$(ijkl)(x) = R_x^h(e_i, e_j, e_k, e_l)$

We will mimic the structure of [Chr1869] as closely as possible; to show our progress through the article the relevant section numbers have been included and equations will frequently be compared to those in [Chr1869].

1., 2.

First of all, our *goal* will be, given two metrics g and h on open subsets U and V of spaces A and B respectively, to determine whether or not there exists a change of variables (diffeomorphism) f from U to V which transforms the metric g into h.

We will start by investigating the necessary conditions for the existence of such an f.

Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $g : U \times A \times A \to \mathbb{K}$, $h : V \times B \times B \to \mathbb{K}$ be metrics on A and B respectively.

Suppose there exists an $f : U \to V$, $f \in C^1(U, B)$ such that for all $a \in U$, $u, v \in A$ we have

$$g_a(u, v) = h_{f(a)}(D_af(u), D_af(v)), \tag{6.2}$$

which corresponds to equation (1.) in [Chr1869].

Suppose $g \in C^1(U)$, $h \in C^1(V)$, and $f \in C^2(U,B)$. Let $a \in U$, $u,v,w \in A$, then from Equation (5.4) and Equation (6.2) we obtain

$$D_a g(u)(v,w) = D_{f(a)}h(D_a f(u))(D_a f(v), D_a f(w)) \\ + h_{f(a)}(D_a^2 f(v,u), D_a f(w)) + h_{f(a)}(D_a f(v), D^2 f(w,u)). \quad (6.3)$$

Following [Chr1869] we will now consider permutations of u, v and w in order to isolate a single $D_a^2 f(u,v)$ term. In light of equation (4.), we define the 3-tensor $B^g \in C^0(U \times A \times A \times A, \mathbb{K})$ by

$$B_a^g(u,v,w) := \frac{1}{2}\Big(D_a g(u)(v,w) - D_a g(w)(u,v) + D_a g(v)(w,u)\Big) \quad (6.4)$$

for all $a \in U$, $u,v,w \in A$. [1] [2]

Note that because of g being symmetric, we obtain (compare with (5.)):

$$B_a^g(u,v,w) = B^g(v,u,w), \qquad D_a g(u)(v,w) = B_a^g(u,v,w) + B_a^g(u,w,v). \quad (6.5)$$

Then by Equation (6.3) and Equation (6.4) we find using symmetry of h and Theorem (5.1.16)

$$\begin{aligned} B_a^g(u,v,w) &= B_{f(a)}^h(D_a f(u), D_a f(v), D_a f(w)) \\ &\quad + \frac{1}{2}\Big(h_{f(a)}(D_a^2 f(v,u), D_a f(w)) + h_{f(a)}(D_a f(v), D_a^2 f(w,u)) \\ &\quad - h_{f(a)}(D_a^2 f(u,w), D_a f(v)) - h_{f(a)}(D_a f(u), D_a^2 f(v,w)) \\ &\quad + h_{f(a)}(D_a^2 f(w,v), D_a f(u)) + h_{f(a)}(D_a f(w), D_a^2 f(u,v))\Big) \\ &= B_{f(a)}^h(D_a f(u), D_a f(v), D_a f(w)) + h_{f(a)}(D_a^2 f(u,v), D_a f(w)). \end{aligned}$$

Hence (compare with (6.) from [Chr1869]) we find for all $a \in U$, $u,v,w \in A$

$$h_{f(a)}(D_a^2 f(u,v), D_a f(w)) + B_{f(a)}^h(D_a f(u), D_a f(v), D_a f(w)) \\ = B_a^g(u,v,w). \quad (6.6)$$

Now we can use the non-degeneracy of g and h to solve this equation, provided that $D_a f : A \to B$ is bijective for all $a \in U$ with a continuous inverse. [3] [4]

[1] This is the first example of an object that we will associate with just the metric g (the definition of B^g depends solely on g itself). We implicitly make the same definition of such an object for the metric h (i.e. a 3-tensor B_a^h defined by Equation (6.4) but with g replaced by h).

[2] One may argue that B^g is 'not a tensor', because B^g does not satisfy the proper transformation relations. With regard to this it should be noted that the definition of a 'tensor' as per Definition (6.1.1) is just a convenient name for a family of multilinear maps with values in $\mathbb{K}$ and therefore this does *not* agree with the usual definition of a tensor as a multilinear object that behaves properly under coordinate transformations. We see however, that the metric (Equation (6.2)), the curvature tensor (Equation (6.12)), and covariant derivatives of a properly transforming covariant tensor (Theorem (6.2.2)) all transform covariantly, as should be expected.

[3] In order to uniquely determine $D_a^2 f$ with Equation (6.6), we need to consider $h_{f(a)}(D_a^2 f(u,v), x)$ for *all* $x \in B$, which means that $D_a f$ should be surjective. On the other hand, we then need to express w in terms of $D_a f(w)$: to be able to do this in a unique way, $D_a f$ should also be injective and because we can only take the inverse of elements of A' and B' (which are continuous), the inverse of $D_a f$ should be continuous.

[4] Note that if A and B satisfy the conditions of Corollary (4.4.6), the inverse of $D_a f$ is continuous whenever $D_a f$ is bijective.

Suppose $D_a f : A \to B$ is bijective for all $a \in U$ and $D_a f^{-1} : B \to A$ ©. Using bijectivity of $D_a f$ we find from Equation (6.6) that

$$h_{f(a)}(D_a^2 f(u,v), x) = B_a^g(u,v,D_a f^{-1}(x)) - B_{f(a)}^h(D_a f(u), D_a f(v), x)$$

for all $a \in U$, $u,v \in A$, $x \in B$.

Hence, using non-degeneracy of h and continuity of $D_a f^{-1}$, $D_a^2 f(u,v)$ is uniquely determined as

$$\begin{aligned} D_a^2 f(u,v) &= \hat{h}_{f(a)}^{-1}(x \mapsto B_a^g(u,v,D_a f^{-1}(x)) - B_{f(a)}^h(D_a f(u), D_a f(v), x)) \\ &= \hat{h}_{f(a)}^{-1}(x \mapsto B_a^g(u,v,D_a f^{-1}(x))) \\ &\quad - \hat{h}_{f(a)}^{-1}(x \mapsto B_{f(a)}^h(D_a f(u), D_a f(v), x)) \end{aligned}$$

for all $a \in U$, $u,v \in A$, and this equation is equivalent to Equation (6.6) by the non-degeneracy of h.

We therefore introduce (compare with (7.)) a map $\Gamma^g : U \times A \times A \to A$ by

$$\Gamma_a^g(u,v) := \hat{g}_a^{-1}\Big(w \mapsto B_a^g(u,v,w)\Big). \tag{6.7}$$

By Definition (6.1.2) and Lemma (6.1.3), $\Gamma^g \in C^0(U \times A \times A, A)$ is a family of 2-linear maps. Note that if $g \in C^k(U)$, then $\Gamma^g \in C^{k-1}(U \times A \times A, A)$. Furthermore, Equation (6.5) implies that for all $a \in U$, $u,v \in A$ we have (compare with (8.))

$$\Gamma_a^g(u,v) = \Gamma_a^g(v,u). \tag{6.8}$$

Note that by definition

$$\hat{h}_{f(a)}^{-1}(x \mapsto B_{f(a)}^h(D_a f(u), D_a f(v), x)) = \Gamma_{f(a)}^h(D_a f(u), D_a f(v)).$$

For the other term in our expression for $D_a^2 f(u,v)$ we first note that for any $w \in A$ we have by Equation (6.2) that $\hat{g}_a(D_a f^{-1}(x))(w) = g_a(D_a f^{-1}(x), w) = h_{f(a)}(x, D_a f(w))$ and hence

$$\begin{aligned} D_a f^{-1}(x) &= \hat{g}_a^{-1}(w \mapsto h_{f(a)}(x, D_a f(w))) \\ &= \hat{g}_a^{-1}(\hat{h}_{f(a)}(x) \circ D_a f). \end{aligned}$$

This permits us to use Lemma (6.1.3) for symmetry of g and after that symmetry of h to obtain for $a \in U$, $u,v \in A$ that

$$\begin{aligned} &\hat{h}_{f(a)}^{-1}(x \mapsto B_a^g(u,v,D_a f^{-1}(x))) \\ &= \hat{h}_{f(a)}^{-1}(x \mapsto B_a^g(u,v,\hat{g}_a^{-1}(\hat{h}_{f(a)}(x) \circ D_a f))) \\ &= \hat{h}_{f(a)}^{-1}(x \mapsto (w \mapsto B_a^g(u,v,w))(\hat{g}_a^{-1}(\hat{h}_{f(a)}(x) \circ D_a f))) \\ &= \hat{h}_{f(a)}^{-1}(x \mapsto (\hat{h}_{f(a)}(x) \circ D_a f)(\hat{g}_a^{-1}(w \mapsto B_a^g(u,v,w)))) \\ &= \hat{h}_{f(a)}^{-1}(x \mapsto \hat{h}_{f(a)}(x)(D_a f(\hat{g}_a^{-1}(w \mapsto B_a^g(u,v,w))))) \\ &= D_a f(\Gamma_a^g(u,v)). \end{aligned}$$

Hence (compare with (9.)) we obtain for all $a \in U$, $u,v \in A$ that

$$D_a^2 f(u,v) + \Gamma_{f(a)}^h(D_a f(u), D_a f(v)) = D_a f(\Gamma_a^g(u,v)). \tag{6.9}$$

By applying $h_{f(a)}(\cdot, D_a f(w))$ on both sides, we again obtain Equation (6.6). Hence Equation (6.6) and Equation (6.9) are equivalent.

So Equation (6.2) implies Equation (6.9), or more precisely: if there exists an $f : U \to V$ satisfying Equation (6.2), $f \in C^2(U, B)$, and $D_a f$ is invertible with continuous inverse for all $a \in U$, then this function f necessarily satisfies Equation (6.9).

3.

Suppose conversely that there exists a C^2 diffeomorphism $f : U \to V$ satisfying Equation (6.9). Then we can define (which is exactly why we need f to be a diffeomorphism, see Lemma (5.1.15)) the metric $i : V \times B \times B \to \mathbb{K}$ by (compare with Equation (6.2))

$$i_b(x, y) := g_{f^{-1}(b)}(D_b f^{-1}(x), D_b f^{-1}(y))$$

which satisfies $i \in C^1(V)$, as $g \in C^1(U)$, and f is a C^2 diffeomorphism. Following the same reasoning as before, i should satisfy Equation (6.9). As h satisfies Equation (6.9) by assumption, we find for all $a \in U$, $u, v \in A$

$$\Gamma^i_{f(a)}(D_a f(u), D_a f(v)) = \Gamma^h_{f(a)}(D_a f(u), D_a f(v)),$$

and hence (f and $D_a f$ are bijective)

$$B^i_b(x, y, z) = B^h_b(x, y, z)$$

for all $b \in V$, $x, y, z \in B$.

We therefore find from Equation (6.5) that

$$D_b i(x)(y, z) = D_b h(x)(y, z)$$

for all $b \in V$, $x, y, z \in B$. But then by Theorem (5.1.8) we have that for all $x, y \in B$ the function $V \to \mathbb{K} : b \mapsto i_b(x, y) - h_b(x, y)$ must be constant. Hence if i and h agree in a single point (i.e. if h satisfies Equation (6.2) at a single point), then they must agree at all points. This leads us to the following theorem.

Theorem 6.2.1
Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $g : U \times A \times A \to \mathbb{K}$, $g \in C^1(U)$, $h : V \times B \times B \to \mathbb{K}$, $h \in C^1(V)$ be metrics on A and B respectively.

Let $f : U \to V$ be a C^2 diffeomorphism, then the following statements are equivalent,

- f satisfies Equation (6.2) for all $u, v \in A$ and all $a \in U$,
- f satisfies Equation (6.2) for all $u, v \in A$ and *some* $a \in U$, and f satisfies Equation (6.9) for all $u, v \in A$ and all $a \in U$.

4.

Suppose that $f \in C^3(U, B)$, $g \in C^2(U)$, and $h \in C^2(V)$ satisfy Equation (6.2) and Equation (6.9), then we can take another derivative of Equation (6.9) using

Equation (5.4), to obtain

$$\begin{aligned}
&D_a^3 f(u,v,w) + D_{f(a)}\Gamma^h(D_a f(w))(D_a f(u), D_a f(v))\\
&+ \Gamma^h_{f(a)}(D_a^2 f(u,w), D_a f(v)) + \Gamma^h_{f(a)}(D_a f(u), D_a^2 f(v,w))\\
&= D_a^2 f(\Gamma^g_a(u,v), w) + D_a f(D_a\Gamma^g(w)(u,v))
\end{aligned}$$

for all $a \in U$, $u, v, w \in A$.

Now swap v and w and subtract this equation from itself (the third order derivatives of f will cancel because of Theorem (5.1.16)) to obtain

$$\begin{aligned}
&0 + D_{f(a)}\Gamma^h(D_a f(w))(D_a f(u), D_a f(v)) - D_{f(a)}\Gamma^h(D_a f(v))(D_a f(u), D_a f(w))\\
&+ \Gamma^h_{f(a)}(D_a^2 f(u,w), D_a f(v)) - \Gamma^h_{f(a)}(D_a^2 f(u,v), D_a f(w)) + 0\\
&= D_a^2 f(\Gamma^g_a(u,v), w) - D_a^2 f(\Gamma^g_a(u,w), v)\\
&\quad + D_a f(D_a\Gamma^g(w)(u,v)) - D_a f(D_a\Gamma^g(v)(u,w)).
\end{aligned}$$

To get rid of the second order derivatives of f we use Equation (6.9):

$$\begin{aligned}
&D_{f(a)}\Gamma^h(D_a f(w))(D_a f(u), D_a f(v)) - D_{f(a)}\Gamma^h(D_a f(v))(D_a f(u), D_a f(w))\\
&+ \Gamma^h_{f(a)}(D_a f(\Gamma^g_a(u,w)), D_a f(v)) - \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(w)), D_a f(v))\\
&- \Gamma^h_{f(a)}(D_a f(\Gamma^g_a(u,v)), D_a f(w)) + \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(v)), D_a f(w)) + 0\\
&= D_a f(\Gamma^g_a(\Gamma^g_a(u,v), w)) - \Gamma^h_{f(a)}(D_a f(\Gamma^g_a(u,v)), D_a f(w))\\
&\quad - D_a f(\Gamma^g_a(\Gamma^g_a(u,w), v)) + \Gamma^h_{f(a)}(D_a f(\Gamma^g_a(u,w)), D_a f(v))\\
&\quad + D_a f(D_a\Gamma^g(w)(u,v)) - D_a f(D_a\Gamma^g(v)(u,w)).
\end{aligned}$$

This can be rearranged to (compare with (12.))

$$\begin{aligned}
&D_{f(a)}\Gamma^h(D_a f(w))(D_a f(u), D_a f(v)) - D_{f(a)}\Gamma^h(D_a f(v))(D_a f(u), D_a f(w))\\
&+ \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(v)), D_a f(w))\\
&- \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(w)), D_a f(v)) + 0\\
&= 0 + D_a f(D_a\Gamma^g(w)(u,v) - D_a\Gamma^g(v)(u,w) + \Gamma^g_a(\Gamma^g_a(u,v), w)\\
&\quad - \Gamma^g_a(\Gamma^g_a(u,w), v)).
\end{aligned} \tag{6.10}$$

5.

We can turn the right-hand side into a 4-tensor by applying $g_a(\cdot, x)$ inside of $D_a f(\cdot)$. In anticipation of this we will rewrite the derivatives $D_a\Gamma^g$, note that with Equation (5.4)

$$\begin{aligned}
D_a B^g(w)(u,v,x) &= D_a(a' \mapsto g_{a'}(\Gamma^g_{a'}(u,v), x))(w)\\
&= D_a g(w)(\Gamma^g_a(u,v), x) + g_a(D_a\Gamma^g(w)(u,v), x),
\end{aligned}$$

so using Equation (6.5) we find

$$\begin{aligned}
g_a(D_a\Gamma^g(w)(u,v), x) &= D_a B^g(w)(u,v,x)\\
&\quad - B^g_a(\Gamma^g_a(u,v), w, x) - B^g_a(x, w, \Gamma^g_a(u,v)).
\end{aligned}$$

Now (Equation (6.4))

$$D_aB^g(w)(u,v,x) = \frac{1}{2}\Big(D_a^2g(u,w)(v,x) - D_a^2g(x,w)(u,v) + D_a^2g(v,w)(x,u)\Big).$$

Therefore (in the last step use Equation (6.7))

$$\begin{aligned}
&g_a(D_a\Gamma^g(w)(u,v) - D_a\Gamma^g(v)(u,w) + \Gamma_a^g(\Gamma_a^g(u,v),w) - \Gamma_a^g(\Gamma_a^g(u,w),v),x)\\
&= \frac{1}{2}\Big(D_a^2g(u,w)(v,x) - D_a^2g(x,w)(u,v) + D_a^2g(v,w)(x,u)\\
&\quad - D_a^2g(u,v)(w,x) + D_a^2g(x,v)(u,w) - D_a^2g(w,v)(x,u)\Big)\\
&\quad - B_a^g(\Gamma_a^g(u,v),w,x) - B_a^g(x,w,\Gamma_a^g(u,v))\\
&\quad + B_a^g(\Gamma_a^g(u,w),v,x) + B_a^g(x,v,\Gamma_a^g(u,w))\\
&\quad + B_a^g(\Gamma_a^g(u,v),w,x) - B_a^g(\Gamma_a^g(u,w),v,x)\\
&= \frac{1}{2}\Big(D_a^2g(u,w)(v,x) - D_a^2g(x,w)(u,v) - D_a^2g(u,v)(w,x) + D_a^2g(x,v)(u,w)\Big)\\
&\quad + g_a(\Gamma_a^g(x,v),\Gamma_a^g(u,w)) - g_a(\Gamma_a^g(x,w),\Gamma_a^g(u,v)).
\end{aligned}$$

Because of this we define the 4-tensor (compare with (14.)) $R^g : U \times A \times A \times A \times A \to \mathbb{K}$ by

$$\begin{aligned}
R_a^g(u,v,w,x) := \frac{1}{2}\Big(&D_a^2g(u,x)(v,w) + D_a^2g(v,w)(u,x)\\
&- D_a^2g(u,w)(v,x) - D_a^2g(v,x)(u,w)\Big)\\
&+ g_a(\Gamma_a^g(u,x),\Gamma_a^g(v,w)) - g_a(\Gamma_a^g(u,w),\Gamma_a^g(v,x))
\end{aligned} \tag{6.11}$$

for all $a \in U$, $u,v,w,x \in A$. By the above (use symmetry of g, Γ and Theorem (5.1.16)):

$$\begin{aligned}
&R_a^g(u,v,w,x)\\
&= \frac{1}{2}\Big(D_a^2g(u,x)(w,v) - D_a^2g(v,x)(u,w) - D_a^2g(u,w)(x,v) + D_a^2g(v,w)(u,x)\Big)\\
&\quad + g_a(\Gamma_a^g(v,w),\Gamma_a^g(u,x)) - g_a(\Gamma_a^g(v,x),\Gamma_a^g(u,w))\\
&= g_a(D_a\Gamma^g(x)(u,w) - D_a\Gamma^g(w)(u,x) + \Gamma_a^g(\Gamma_a^g(u,w),x) - \Gamma_a^g(\Gamma_a^g(u,x),w),v)
\end{aligned}$$

Now apply $h_{f(a)}(\cdot, D_a f(x))$ on both sides of Equation (6.10) to obtain

$$
\begin{aligned}
&R^h_{f(a)}(D_a f(u), D_a f(x), D_a f(v), D_a f(w))\\
&= h_{f(a)}\Big(D_{f(a)}\Gamma^h(D_a f(w))(D_a f(u), D_a f(v))\\
&\quad - D_{f(a)}\Gamma^h(D_a f(v))(D_a f(u), D_a f(w))\\
&\quad + \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(v)), D_a f(w))\\
&\quad - \Gamma^h_{f(a)}(\Gamma^h_{f(a)}(D_a f(u), D_a f(w)), D_a f(v)), D_a f(x)\Big)\\
&= h_{f(a)}\Big(D_a f(D_a\Gamma^g(w)(u,v) - D_a\Gamma^g(v)(u,w) + \Gamma^g_a(\Gamma^g_a(u,v), w)\\
&\quad - \Gamma^g_a(\Gamma^g_a(u,w), v)), D_a f(x)\Big)\\
&\overset{(6.2)}{=} g_a\Big(D_a\Gamma^g(w)(u,v) - D_a\Gamma^g(v)(u,w) + \Gamma^g_a(\Gamma^g_a(u,v), w)\\
&\quad - \Gamma^g_a(\Gamma^g_a(u,w), v)), x\Big)\\
&= R^g_a(u, x, v, w).
\end{aligned}
$$

Hence (compare with (15.)) we find for all $a \in U$, $u, v, w, x \in A$ that

$$R^g_a(u, v, w, x) = R^h_{f(a)}(D_a f(u), D_a f(v), D_a f(w), D_a f(x)). \tag{6.12}$$

Furthermore, from Equation (6.11) and symmetry of g, Γ, and higher order derivatives we obtain (compare with (16.)) for all $a \in U$, $u, v, w, x \in A$ that

$$
\begin{aligned}
R^g_a(u, v, w, x) &= -R^g_a(v, u, w, x),\\
R^g_a(u, v, w, x) &= -R^g_a(u, v, x, w),\\
R^g_a(u, v, w, x) &= R^g_a(w, x, u, v),\\
R^g_a(u, v, w, x) &= -R^g_a(u, w, x, v) - R^g_a(u, x, v, w).
\end{aligned} \tag{6.13}
$$

6.

Considering Equation (6.2) and Equation (6.12) we see that it might be rewarding to consider a general k-tensor that satisfies a similar transformation rule.

So let $E^A : U \times A \times \ldots \times A \to \mathbb{K}$, $E^B : V \times B \times \ldots \times B \to \mathbb{K}$ be k-tensors for some $k \in \mathbb{N}$ on A and B respectively. Furthermore suppose that $E^A \in C^1(U)$, $E^B \in C^1(V)$ and that they satisfy for all $a \in U$, $u_1, \ldots, u_k \in A$ the equation

$$E^A_a(u_1, \ldots, u_k) = E^B_{f(a)}(D_a f(u_1), \ldots, D_a f(u_k)). \tag{6.14}$$

Then we can in the same way as we did for g and h take the derivative of this equation using Equation (5.4), and use Equation (6.9) to obtain

$$
\begin{aligned}
&D_a E^A(v)(u_1,\dots,u_k)\\
&= D_{f(a)}E^B(D_af(v))(D_af(u_1),\dots,D_af(u_k))\\
&\quad + E^B_{f(a)}(D_a^2f(u_1,v),\dots,D_af(u_k)) + \dots + E^B_{f(a)}(D_af(u_1),\dots,D_a^2f(u_k,v))\\
&\overset{(6.9)}{=} D_{f(a)}E^B(D_af(v))(D_af(u_1),\dots,D_af(u_k))\\
&\quad - \Big(E^B_{f(a)}(\Gamma^h_{f(a)}(D_af(u_1),D_af(v)),\dots,D_af(u_k))\\
&\qquad + \dots + E^B_{f(a)}(D_af(u_1),\dots,\Gamma^h_{f(a)}(D_af(u_k),D_af(v)))\Big)\\
&\quad + \Big(E^B_{f(a)}(D_af(\Gamma^g_a(u_1,v)),\dots,D_af(u_k))\\
&\qquad + \dots + E^B_{f(a)}(D_af(u_1),\dots,D_af(\Gamma^g_a(u_k,v)))\Big)\\
&\overset{(6.14)}{=} D_{f(a)}E^B(D_af(v))(D_af(u_1),\dots,D_af(u_k))\\
&\quad - \Big(E^B_{f(a)}(\Gamma^h_{f(a)}(D_af(u_1),D_af(v)),\dots,D_af(u_k))\\
&\qquad + \dots + E^B_{f(a)}(D_af(u_1),\dots,\Gamma^h_{f(a)}(D_af(u_k),D_af(v)))\Big)\\
&\quad + \Big(E^A_a(\Gamma^g_a(u_1,v),\dots,u_k) + \dots + E^A_a(u_1,\dots,\Gamma^g_a(u_k,v))\Big).
\end{aligned}
$$

Define the $(k+1)$-tensor $\nabla E^A : U \times A \times A \times \dots \times A \to \mathbb{K}$ by

$$
\begin{aligned}
\nabla E^A_a(u_0,u_1,\dots,u_k) &:= D_aE^A(u_0)(u_1,\dots,u_k)\\
&\quad - \Big(E^A_a(\Gamma^g_a(u_0,u_1),\dots,u_k) + \dots + E^A_a(u_1,\dots,\Gamma^g_a(u_0,u_k))\Big)
\end{aligned}
\tag{6.15}
$$

for all $a \in U$, $u_0,u_1,\dots,u_k \in A$. Then by the above for $v = u_0$ this new tensor satisfies

$$
\nabla E^A_a(u_0,u_1,\dots,u_k) = \nabla E^B_{f(a)}(D_af(u_0),D_af(u_1),\dots,D_af(u_k))
$$

and hence transforms in exactly the same way as E^A and E^B: it satisfies Equation (6.14) for $k+1$.

Therefore we now arrive at the following statement.

Theorem 6.2.2
Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$ be a C^2 diffeomorphism, and $g : U \times A \times A \to \mathbb{K}$, $g \in C^1(U)$, $h : V \times B \times B \to \mathbb{K}$, $h \in C^1(V)$ be metrics on A and B respectively. Suppose f, g and h satisfy Equation (6.2) for all $u, v \in A$ and $a \in U$.

Let $k \in \mathbb{N}$, $E^A : U \times A \times \dots \times A \to \mathbb{K}$, $E^B : V \times B \times \dots \times B \to \mathbb{K}$ be k-tensors on A and B respectively that are both C^1.

If E^A and E^B satisfy Equation (6.14) for all $a \in U$, then the $(k+1)$-tensors ∇E^A and ∇E^B defined by Equation (6.15) on A and B respectively satisfy Equation (6.14) for $k+1$ and all $a \in U$.

So we can 'take the derivative' [5] of any k-tensor satisfying Equation (6.14) to obtain a $(k+1)$-tensor which also satisfies Equation (6.14).

[5] This really is the *covariant derivative* with respect to the Levi-Civita connection induced by the metric g.

Taking a look at this procedure for $k = 2$ with $E^A = g$ and $E^B = h$ we find using Equation (6.5) that

$$\begin{aligned}\nabla g_a(u,v,w) &= D_a g(u)(v,w) - g_a(\Gamma^g_a(u,v),w) - g_a(v,\Gamma^g_a(u,w))\\ &= D_a g(u)(v,w) - B^g_a(u,v,w) - B^g_a(u,w,v)\\ &= D_a g(u)(v,w) - D_a g(u)(v,w) = 0.\end{aligned}$$

So $\nabla g = 0$, and similarly $\nabla h = 0$.[6] Therefore this derivative does not give us any new equations directly from Equation (6.2).

7., 8.

Suppose $f \in C^3(U,B)$, $g \in C^1(U)$, $h \in C^1(V)$, and that f, g and h satisfy Equation (6.2) and Equation (6.9) for all $a \in U$. Then they satisfy Equation (6.14) for $k = 2$ with $E^A = g$ and $E^B = h$. Using Theorem (6.2.2) we can consider Equation (6.14) for $k = 3$ with ∇g and ∇h, but we already saw that $\nabla g = \nabla h = 0$ do not give any new equations. Looking at Equation (6.12) however we find that Equation (6.2) implies, provided $g \in C^2(U)$ and $h \in C^2(V)$, that we also satisfy Equation (6.14) for $k = 4$ with $E^A = R^g$ and $E^B = R^h$. Now we can, provided $g \in C^3(U)$ and $h \in C^3(V)$, again use Theorem (6.2.2) and find that we satisfy Equation (6.14) for $k = 5$ with $E^A = \nabla R^g$, $E^B = \nabla R^h$. Applying Theorem (6.2.2) again we satisfy Equation (6.14) for $k = 6$ by considering $\nabla(\nabla R^g)$ and $\nabla(\nabla R^h)$, and we can continue this way indefinitely if $g \in C^\infty(U)$ and $h \in C^\infty(V)$.

Theorem 6.2.3
Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$ be a C^3 diffeomorphism, and $g : U \times A \times A \to \mathbb{K}$, $g \in C^1(U)$, $h : V \times B \times B \to \mathbb{K}$, $h \in C^1(V)$ be metrics on A and B respectively.

Suppose f, g and h satisfy Equation (6.2) for all $u, v \in A$ and $a \in U$ and that there is an $l \in \mathbb{N}$, $l \geq 2$ such that $g \in C^l(U)$ and $h \in C^l(V)$.

Then we obtain a chain of instances of Equation (6.14), each of which implies the next (via Equation (6.12) for $k = 2$ and via Theorem (6.2.2) for $k \geq 4$, write $\nabla^m R^g := \underbrace{\nabla \ldots \nabla}_{m} R^g$):

k	E^A	E^B
2	g	h
4	R^g	R^h
5	∇R^g	∇R^h
6	$\nabla^2 R^g$	$\nabla^2 R^h$
$\ldots$	$\ldots$	$\ldots$
$l+2$	$\nabla^{l-2} R^g$	$\nabla^{l-2} R^h$.

6.3 Digression

This is the point where we will diverge from [Chr1869] because of our more general (typically infinite-dimensional) setting, which is not compatible with

[6] Again in a more modern context: the metric is covariantly constant.

the theory of invariants and equation counting that is used in Sections 9. to 12. of the article. [7]

We will first introduce two very useful definitions which provide a pleasant shorthand for Equation (6.14) and permit us to push and pull k-tensors from one open set to another using the diffeomorphisms between these open sets.

Definition 6.3.1: Pullback

Let A, B /$\mathbb{K}$, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$, $f \in C^1(U, B)$, $k \in \mathbb{N}$, and $g : V \times B^k \to \mathbb{K}$ a k-tensor on B.

Then the *pullback of g by f* is defined as the k-tensor $f^*g : U \times A^k \to \mathbb{K}$ on A, for all $a \in U$, $u_1, \ldots, u_k \in A$ given by

$$(f^*g)_a(a_1, \ldots, a_k) := g_{f(a)}(D_a f(u_1), \ldots, D_a f(u_k)). \tag{6.16}$$

Definition 6.3.2: Pushforward

Let A, B /$\mathbb{K}$, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$ a C^1 diffeomorphism, $k \in \mathbb{N}$, and $g : U \times A^k \to \mathbb{K}$ a k-tensor on A.

Then the *pushforward of g by f* is defined as the k-tensor on B, $f_*g : V \times B^k \to \mathbb{K}$, for all $b \in V$, $v_1, \ldots, v_k \in B$ given by

$$(f_*g)_b(v_1, \ldots, v_k) := g_{f^{-1}(b)}([D_{f^{-1}(b)}f]^{-1}(v_1), \ldots, [D_{f^{-1}(b)}f]^{-1}(v_k)). \tag{6.17}$$

Lemma 6.3.3

Let A, B /$\mathbb{K}$, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$, $f \in C^1(U, B)$, $k \in \mathbb{N}$, and $g : V \times B^k \to \mathbb{K}$ a k-tensor on B.

If for $l \in \mathbb{N}$, $f \in C^{l+1}(U, B)$ and $g \in C^l(V)$, then $(f^*g) \in C^l(U)$.

Suppose f is a C^1 diffeomorphism and let $h : U \times A^k \to \mathbb{K}$ be a k-tensor on A, then

$$f_*h = (f^{-1})^*h.$$

In particular, if f is a C^{l+1} diffeomorphism and $h \in C^l(U)$, then $f_*h \in C^l(V)$.

Proof. Suppose $f \in C^{l+1}(U, V)$ and $g \in C^l(V)$. Then using induction, the expression for f^*g from Definition (6.3.1) and Equation (5.4) we see that $(f^*g) \in C^l(U)$.

Suppose f is a C^1 diffeomorphism, then using Lemma (5.1.15) we find that for all $b \in V$, $v_1, \ldots, v_k \in B$ that

$$\begin{aligned}
(f_*h)_b(v_1, \ldots, v_k) &= h_{f^{-1}(b)}([D_{f^{-1}(b)}f]^{-1}(v_1), \ldots, [D_{f^{-1}(b)}f]^{-1}(v_k)) \\
&= h_{f^{-1}(b)}(D_b f^{-1}(v_1), \ldots, D_b f^{-1}(v_k)) \\
&= ((f^{-1})^*h)_b(v_1, \ldots, v_k),
\end{aligned}$$

so $f_*h = (f^{-1})^*h$. Applying the first part of the lemma we therefore see that if $f^{-1} \in C^{l+1}(U, B)$ and $h \in C^l(U)$, then $f_*h = (f^{-1})^*h \in C^l(V)$. □

Lemma 6.3.4

Let A, B, C /$\mathbb{K}$, $U \subseteq A$, $V \subseteq B$, $W \subseteq C$ open. Let $f : U \to V$, $f \in C^1(U, B)$, and $g : V \to W$, $g \in C^1(V, C)$.

[7] It should be noted that Christoffel does not give any explicit means or examples of actually obtaining these invariants in [Chr1869].

Then for any k-tensor $h : W \times C^k \to \mathbb{K}$ on C we have

$$(g \circ f)^* h = f^*(g^* h). \tag{6.18}$$

If f and g are C^1 diffeomorphisms, then for any k-tensor $i : U \times A^k \to \mathbb{K}$ on A

$$(g \circ f)_* i = g_*(f_* i). \tag{6.19}$$

In particular if f is a C^1 diffeomorphism, h a k-tensor on B, and i a k-tensor on A

$$f_*(f^* h) = h, \qquad f^*(f_* i) = i.$$

Proof. Let $a \in U$, $u_1, \dots, u_k \in A$, then using Theorem (5.1.8)

$$\begin{aligned}((g \circ f)^* h)_a(u_1, \dots, u_k) &= h_{(g\circ f)(a)}(D_a(g \circ f)(u_1), \dots, D_a(g \circ f)(u_k)) \\ &= h_{g(f(a))}(D_{f(a)} g(D_a f(u_1)), \dots, D_{f(a)} g(D_a f(u_k))) \\ &= (g^* h)_{f(a)}(D_a f(u_1), \dots, D_a f(u_k)) \\ &= (f^*(g^* h))_a(u_1, \dots, u_k),\end{aligned}$$

which shows that $(g \circ f)^* h = f^*(g^* h)$.

Using Lemma (6.3.3) we see that if f and g are C^1 diffeomorphisms, then $(g \circ f)_* i = ((g \circ f)^{-1})^* i = (f^{-1} \circ g^{-1})^* i = (g^{-1})^*((f^{-1})^* i) = g_*(f_* i)$.

Using the same trick we also see that $f_*(f^* h) = (f^{-1})^*(f^* h) = (f \circ f^{-1})^* h = (\mathrm{id}_V)^* h = h$ and similarly $f^*(f_* i) = (\mathrm{id}_U)^* i = i$. □

Lemma 6.3.5

Let A, B Vs/$\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $f : U \to V$, be a C^1 diffeomorphism, and $h : V \times B \times B \to \mathbb{K}$ a metric on B.

Then the pullback $f^* h$ is a metric on A.

Proof. We already know that $g := f^* h : U \times A \times A \to \mathbb{K}$ is a 2-tensor. As h is symmetric, it is immediate from Equation (6.16) that g is symmetric. Now for $a \in U$, $u, v \in A$ we have $(\hat{g}_a(u))(v) = g_a(u, v) = h_{f(a)}(D_a f(u), D_a f(v)) = (\hat{h}_{f(a)}(D_a f(u)))(D_a f(v)) = (\hat{h}_{f(a)}(D_a f(u)) \circ D_a f)(v)$. So for $a \in U$, $u \in A$

$$\hat{g}_a(u) = \hat{h}_{f(a)}(D_a f(u)) \circ D_a f$$

from which we find that $\hat{g}_a$ is bijective with inverse

$$\hat{g}_a^{-1}(i) = [D_a f]^{-1}(\hat{h}_{f(a)}^{-1}(i \circ [D_a f]^{-1}))$$

for all $i \in A'$. Furthermore, as $\hat{h}^{-1}$ C and $f^{-1} \in C^1(U, B)$, we see that $\hat{g}^{-1}$ C by Lemma (5.1.15). Therefore g is a metric. □

Because we will soon be using multiple metrics and diffeomorphisms simultaneously, we will from now on denote the diffeomorphisms by greek characters ϕ, χ, ψ, ... to avoid confusion.

Theorem 6.3.6

[8] Let A, B, C, D Vs/$\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$, $W \subseteq C$, $X \subseteq D$ open. Let

[8]This theorem is nothing but an expression of the fact that in the setting of Riemannian manifolds we can look at the metrics g and h in coordinate charts of our choosing.

$g : U \times A \times A \to \mathbb{K}$, $h : V \times B \times B \to \mathbb{K}$ be metrics on A and B respectively. Let for $k \in \mathbb{N}$, $\phi : W \to U$, $\chi : X \to V$ be C^k diffeomorphisms.

Then there exists a C^k diffeomorphism $\psi : U \to V$ such that $g = \psi^* h$ if and only if there exists a C^k diffeomorphism $\omega : W \to X$ such that $(\phi^* g) = \omega^*(\chi^* h)$,

$$\begin{tikzcd} \phi^* g & W \arrow[r, "\omega"] \arrow[d, "\phi"'] & X \arrow[d, "\chi"] & \chi^* h \\ g & U \arrow[r, "\psi"'] & V & h. \end{tikzcd}$$

Proof. Suppose there exists a C^k diffeomorphism $\psi : U \to V$ such that $g = \psi^* h$. Then $\omega := \chi^{-1} \circ \psi \circ \phi : W \to X$ is a C^k diffeomorphism, and by Equation (6.18) $\omega^*(\chi^* h) = (\chi \circ \omega)^* h = (\chi \circ \chi^{-1} \circ \psi \circ \phi)^* h = \phi^*(\psi^* h) = \phi^* g$.

Suppose conversely that there exists a C^k diffeomorphism $\omega : W \to X$ such that $\phi^* g = \omega^*(\chi^* h)$. Choose $\psi := \chi \circ \omega \circ \phi^{-1} : U \to V$ which is a C^k diffeomorphism, then by Equation (6.18) $\psi^* h = (\chi \circ \omega \circ \phi^{-1})^* h = (\phi^{-1})^*(\omega^*(\chi^* h)) = (\phi^{-1})^*(\phi^* g) = (\phi \circ \phi^{-1})^* g = g$. □

Although it may not immediately be apparent, this theorem permits us to simplify the question we asked at Equation (6.2). From our previous considerations (being Theorem (6.2.1), Theorem (6.2.2), and Theorem (6.2.3)) we see that if we want to compare g and h in a way compatible with [Chr1869], we need U and V to be at least C^3 diffeomorphic (via the map f).

Corollary 6.3.7

Let A, B Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$, $V \subseteq B$ open. Let $g : U \times A \times A \to \mathbb{K}$, $h : V \times B \times B \to \mathbb{K}$ be metrics on A and B respectively.

Suppose for $k \in \mathbb{N}$ that U and V are C^k diffeomorphic via some C^k diffeomorphism $\chi : U \to V$.

Then there exists a C^k diffeomorphism $f : U \to V$ such that f, g, and h satisfy Equation (6.2) if and only if there exists a C^k diffeomorphism $\omega : U \to U$ such that $g = \omega^*(\chi^* h)$.

Proof. Simply apply Theorem (6.3.6) for $C = D = A$, $W = X = U$, $\phi = \mathrm{id}_U$, and $f = \psi$. □

This permits us to reduce without loss of generality to the case where $B = A$ and $V = U$ by considering the metric (Lemma (6.3.5)) $\chi^* h$ instead of h. Using Lemma (6.3.4) we see that for a diffeomorphism $\phi : U \to U$, $g = \phi^* h$ is equivalent to $\phi_* g = h$, so we can use pushforwards just as well as pullbacks.

Because we can assume $V = U$, the collection of all diffeomorphisms that may take h into g or vice versa can be composed with each other and therefore they form a group.

Definition 6.3.8: Diffeomorphism group

Let A Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open, and $k \in \mathbb{N}$.

Then the *k-diffeomorphism group of U* is defined to be

$$\mathcal{G}^k_U := \{\phi : U \to U |\ \phi \text{ is a } C^k \text{ diffeomorphism } \}$$

together with identity $\mathrm{id}_U : U \to U : a \mapsto a$, multiplication $(\phi, \chi) \mapsto \phi \circ \chi$, and inversion $\phi \mapsto \phi^{-1}$.

The *collection of C^k metrics on U* is defined to be

$$\mathcal{M}_U^k := \{g : U \times A \times A \to \mathbb{K} |\ g \text{ is a metric, } g \in C^k(U)\ \}.$$

By the above, $\mathcal{G}_U^k$ has a natural action on $\mathcal{M}_U^l$ for all $k, l \in \mathbb{N}$, $k > l$ given by the pushforward

$$\mathcal{G}_U^k \times \mathcal{M}_U^l \to \mathcal{M}_U^l : (\phi, g) \mapsto \phi \cdot g := \phi_* g. \tag{6.20}$$

That this indeed is well-defined action can be verified with Lemma (6.3.3), Lemma (6.3.5), and Lemma (6.3.4): $\chi \cdot (\phi \cdot g) = \chi_*(\phi_* g) = (\chi \circ \phi)_* g = (\chi \circ \phi) \cdot g$ and $\mathrm{id}_U \cdot g = (\mathrm{id}_U)_* g = g$.

We see that now the question of finding a C^k diffeomorphism $f : U \to V$ such that f, g, and h satisfy Equation (6.2) is equivalent to $\chi^* h$ lying in the orbit $\mathcal{G}_U^k \cdot g$, where $\chi : U \to V$ is any C^k diffeomorphism. Therefore we can answer our question if we know the orbit of g under the action of $\mathcal{G}_U^k$. Unfortunately, the group $\mathcal{G}_U^k$ is extremely complicated.

We therefore will not concentrate on the entire orbit of g, but on the orbit of the Taylor expansion (recall Theorem (5.3.11)) of g at a fixed point $a \in U$. Using Theorem (6.3.6) and the diffeomorphism $U \to (U - a) : a_1 \mapsto a_1 - a$ we see that we may assume that U is an open neighbourhood of $0 = a$. So we from now on assume that A Vs $/\mathbb{K}$ T2 LC UC (necessary to use Theorem (5.3.11)) and U an open neighbourhood of 0 in A.

Using Theorem (5.1.8), Equation (5.4), and Equation (6.16) we see that for $g \in \mathcal{M}_U^l$, $\phi \in \mathcal{G}_U^k$ with $k > l$ both large enough and $\phi(0) = 0$ we have that the Taylor sequence of g at 0 and the Taylor sequence of $\phi^* g$ at 0 look as Table 6.1 (for $\phi_* g$ simply use $\phi \to \phi^{-1}$, Lemma (6.3.3)). It is clear that for higher order

Order	g	$\phi^* g$
0	$g_0(u, v)$	$g_0(D_0\phi(u), D_0\phi(v))$
1	$D_0 g(w)(u, v)$	$D_0 g(D_0\phi(w))(D_0\phi(u), D_0\phi(v))$ $+g_0(D_0^2\phi(u, w), D_0\phi(v)) + g_0(D_0\phi(u), D_0^2\phi(v, w))$
2	$D_0^2 g(w, x)(u, v)$	$D_0^2 g(D_0\phi(w), D_0\phi(x))(D_0\phi(u), D_0\phi(v))$ $+D_0 g(D_0\phi(w))(D_0^2\phi(u, x), D_0\phi(v))$ $+D_0 g(D_0\phi(w))(D_0\phi(u), D_0^2\phi(v, x))$ $+D_0 g(D_0\phi(x))(D_0^2\phi(u, w), D_0\phi(v))$ $+g_0(D_0^3\phi(u, w, x), D_0\phi(v)) + g_0(D_0^2\phi(u, w), D_0^2\phi(v, x))$ $+D_0 g(D_0\phi(x))(D_0\phi(u), D_0^2\phi(v, w))$ $+g_0(D_0^2\phi(u, x), D_0^2\phi(v, w)) + g_0(D_0\phi(u), D_0^3\phi(v, w, x))$
$\vdots$	$\vdots$	$\vdots$

Table 6.1: Relation between the Taylor sequence of a metric g and of its pullback $\phi^* g$ by a diffeomorphism ϕ. Obtained by repeatedly using Theorem (5.1.8) and Equation (5.4).

expansions, the transformation rules become extremely convoluted: this is not a very practical approach.

From [Chr1869] we obtained, in the form of Theorem (6.2.3), that for R^g, $\nabla R^g, \ldots$ we have the results in Table 6.2. For these terms, all expressions retain

g	$\phi^* g$
$g_0(u,v)$	$g_0(D_0\phi(u), D_0\phi(v))$
$R^g_0(u,v,w,x)$	$R^g_0(D_0\phi(u), D_0\phi(v), D_0\phi(w), D_0\phi(x))$
$\nabla R^g_0(u,v,w,x,y)$	$\nabla R^g_0(D_0\phi(u), D_0\phi(v), D_0\phi(w), D_0\phi(x), D_0\phi(y))$
$\vdots$	$\vdots$

Table 6.2: Chain of equations from Theorem (6.2.3) for a metric g and its pullback $\phi^* g$ by a diffeomorphism ϕ.

their original form under pullback (apart from an introduced $D_0\phi$).

However, while we know that near 0 the Taylor series $(g_0, D_0 g, D_0^2 g, \ldots)$ gives a reasonable description of g (Corollary (5.3.12)), it is quite unclear whether or not we can approximate g by $(g_0, R^g_0, \nabla R^g_0, \ldots)$. We can determine the sequence $(g_0, R^g_0, \nabla R^g_0, \ldots)$ directly from the Taylor series $(g_0, D_0 g, D_0^2 g, \ldots)$ using Equation (6.1), Equation (6.11), and Equation (6.15). However, it is not clear whether or not we can recover $(g_0, D_0 g, D_0^2 g, \ldots)$ again from $(g_0, R^g_0, \nabla R^g_0, \ldots)$. If we could, this would guarantee via Corollary (5.3.12) that $(g_0, R^g_0, \nabla R^g_0, \ldots)$ describes the metric near 0, therefore we will now investigate whether or not this is possible.

By Theorem (6.3.6) we can start by considering $\phi^* g$ (which lies in the same orbit as g) instead of g, for a diffeomorphism ϕ which simplifies g as much as possible. Then for this simplified g, we will try to determine to what degree we can recover $(g_0, D_0 g, D_0^2 g, \ldots)$ from $(g_0, R^g_0, \nabla R^g_0, \ldots)$. [9]

It turns out (Theorem (6.5.1)) that to actually obtain this 'simplifying diffeomorphism' ϕ we are going to need Theorem (5.5.8) and Theorem (5.5.10) for existence, which requires us to demand that A $\boxed{\text{Ba}}$ $/\mathbb{K}$. [10]

However, we will first discuss what we consider to be a 'simple' metric.

[9] This simplification of g actually seems to be necessary. For general g there is no apparent way in which such a recovery can be made; one can even show that $D_0^2 g(w,x)(u,v)$ *cannot* be recovered from any linear combination of $R^g_0(u,v,w,x)$ where we permute u, v, w, and x (this is immediately clear if A is finite-dimensional: the number of independent components of $D_0^2 g$ with respect to a basis $\{e_1, \ldots, e_k\}$ of A, $(\frac{1}{2}k(k+1))^2$, is strictly greater than that of R^g_0, $\frac{1}{12}k^2(k^2-1)$, in the general case it can be rewritten as a conflicting system of linear equations).

[10] Quite a pity, up until now we have been working in a very general context. However, Example (5.4.3) and Example (5.4.4) show that this is really necessary. Furthermore, A is almost a Hilbert space (so in particular almost $\boxed{\text{Ba}}$) because of the existence of the metric g: necessarily $A \simeq A'$. In particular, if there exists a point a at which $\hat{g}_a$ is negative definite on a finite-dimensional subspace of A and positive semidefinite on the complement, then A is a Hilbert space.

6.4 Simple metrics

We will start by considering the notion of a geodesic: a generalisation of the concept of a straight line, with respect to the metric g. [11] [12] Geodesics will be curves γ for which the 'total kinetic energy' of the curve, given by $\int \left(\alpha \mapsto \frac{1}{2} g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha))\right)$, is extremal with respect to variations of γ which leave the end-points of γ fixed. Theorem (5.3.14) shows us that such variations are completely described by the Lagrange map of the kinetic energy.

Let A /$\mathbb{K}$, and $U \subseteq A$ open. Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A with $g \in C^2(U)$.

Consider the function $G : U \times A \to \mathbb{K}$ given by

$$G(a, u) := \frac{1}{2} g_a(u, u)$$

which is C^2.

Then for any curve $\gamma : S \to U$, $\gamma \in C^2(S, A)$ with $S \subseteq \mathbb{R}$ an open interval, we have for $\alpha \in S$, $u \in A$ (see Theorem (5.3.14) and use Equation (5.4))

$$\begin{aligned}
\operatorname{La}(G, \gamma)(\alpha, u) &= \frac{\partial}{\partial \alpha}\Big(g_{\gamma(\alpha)}(\gamma'(\alpha), u)\Big) - \frac{1}{2} D_{\gamma(\alpha)} g(u)(\gamma'(\alpha), \gamma'(\alpha)) \\
&= D_{\gamma(\alpha)} g(\gamma'(\alpha))(\gamma'(\alpha), u) + g_{\gamma(\alpha)}(\gamma''(\alpha), u) \\
&\quad - \frac{1}{2} D_{\gamma(\alpha)} g(u)(\gamma'(\alpha), \gamma'(\alpha)) \\
&\overset{(6.4)}{=} B^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha), u) + g_{\gamma(\alpha)}(\gamma''(\alpha), u) \\
&\overset{(6.7)}{=} g_{\gamma(\alpha)}\Big(\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha)), u\Big) \\
&= \Big(\hat{g}_{\gamma(\alpha)}\Big(\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha))\Big)\Big)(u).
\end{aligned}$$

Hence $\operatorname{La}(G, \gamma)(\alpha, u) = 0$ for all $\alpha \in S$, $u \in A$ if and only if $\hat{g}_{\gamma(\alpha)}\Big(\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha))\Big) = 0$ for all $\alpha \in S$ if and only if (Definition (6.1.2)) for all $\alpha \in S$ we have

$$\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha)) = 0.$$

Definition 6.4.1: Geodesic

Let A /$\mathbb{K}$, $U \subseteq A$ open. Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A, $g \in C^1(U)$.

Then we call a curve $\gamma : S \to U$, $\gamma \in C^2(S, A)$, with $S \subseteq \mathbb{R}$ an open interval a *geodesic with respect to* g if for all $\alpha \in S$, γ satisfies

$$\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha)) = 0. \tag{6.21}$$

[11] For metrics g that are positive definite, $g_a(u, u) > 0$ for all $u \neq 0$, the geodesics are actually (see Section 10 of [Ban2008]) locally the paths of shortest length as measured by $\int_{\operatorname{dom}\gamma} \sqrt{g_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha))}\, \mathrm{d}\alpha$. One can compare this with straight lines $\gamma(\alpha) = a + \alpha u$ in $\mathbb{R}^k$ which are paths of shortest length with respect to $\int_{\operatorname{dom}\gamma} \sqrt{\langle \gamma'(\alpha), \gamma'(\alpha) \rangle}\, \mathrm{d}\alpha = \int_{\operatorname{dom}\gamma} \|\gamma'(\alpha)\|\, \mathrm{d}\alpha$, which is the usual notion of length.

[12] Another interpretation of geodesics is that of paths of free-falling, that is, not subject to external forces, particles in the theory of general relativity. See Chapter 4 from [Wal1984].

Let $\gamma : S \to U$ be a geodesic with respect to g. Then using Equation (5.4) we find that

$$\begin{aligned}
\frac{\partial}{\partial \alpha}\Big(g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha))\Big) &= D_{\gamma(\alpha)} g(\gamma'(\alpha))(\gamma'(\alpha),\gamma'(\alpha)) \\
&\quad + 2\, g_{\gamma(\alpha)}(\gamma''(\alpha),\gamma'(\alpha)) \\
&\overset{(6.5)}{=} 2\, B^g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha)) + 2\, g_{\gamma(\alpha)}(\gamma''(\alpha),\gamma'(\alpha)) \\
&\overset{(6.7)}{=} 2\, g_{\gamma(\alpha)}\Big(\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha)),\gamma'(\alpha)\Big) \\
&\overset{(6.21)}{=} 0.
\end{aligned}$$

Therefore we find with Theorem (5.1.8) that for a geodesic $\gamma : S \to U$,

$$g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha)) = g_{\gamma(\beta)}(\gamma'(\beta),\gamma'(\beta))$$

for all $\alpha, \beta \in S$. [13]

So we obtain the following lemma.

Lemma 6.4.2

Let A Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open. Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A, $g \in C^2(U)$.

Let $\gamma : S \to U$, $\gamma \in C^2(S, A)$, with $S \subseteq \mathbb{R}$ an open interval. Then γ is a geodesic with respect to g if and only if

$$\mathrm{La}\left(\left(U \times A \to \mathbb{K} : (a,u) \mapsto \frac{1}{2}\, g_a(u,u)\right), \gamma\right)(\alpha, u) = 0 \tag{6.22}$$

for all $\alpha \in S$, $u \in A$.

If this is the case, we have for all $\alpha, \beta \in S$

$$g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha)) = g_{\gamma(\beta)}(\gamma'(\beta),\gamma'(\beta)). \tag{6.23}$$

We now call a metric simple at a certain point if the geodesics, the straight lines with respect to g, emanating from this point, are straight in the usual sense (see Lemma (6.4.4)).

Definition 6.4.3: Simple metric

Let A Vs $/\mathbb{K}$ T2 LC, $U \subseteq A$ open.

Then a metric $g : U \times A \times A \to \mathbb{K}$ on A, $g \in C^1(U)$ is called *a simple metric at* $a \in U$ if there exists a neighbourhood U_1 of 0 in $U - a$ such that for all $u \in U_1$ we have

$$\Gamma^g_{a+u}(u,u) = 0. \tag{6.24}$$

Suppose g is a simple metric at $a \in U$, satisfying Equation (6.24) for all $u \in U_1$ where U_1 is an open abc (A LC) neighbourhood of 0 in $U - a$. Let $u \in U_1$, then, as U is balanced, for all $\alpha \in B_{\mathbb{K}}(0,1)$ we have $\alpha\, u \in U_1$, so $\Gamma^g_{a+\alpha\, u}(\alpha\, u, \alpha\, u) = \alpha^2\, \Gamma^g_{a+\alpha\, u}(u,u) = 0$. Define $\gamma :]-1,1[\to U$ by $\gamma(\alpha) := a + \alpha\, u$, then

$$\begin{aligned}
\gamma''(\alpha) + \Gamma^g_{\gamma(\alpha)}(\gamma'(\alpha),\gamma'(\alpha)) &= 0 + \Gamma^g_{a+\alpha\, u}(u,u) \\
&= 0,
\end{aligned}$$

[13] So the infinitesimal length squared of geodesics with respect to g is constant.

so $\alpha \mapsto a + \alpha\, u$ is a geodesic. Therefore, if g is simple at a, all straight lines emanating from a are geodesics with respect to g.

Suppose conversely that there exists an open abc neighbourhood U_1 of 0 in $U - a$ such that for all $u \in U_1$, the map $]-1,1[\to U : \alpha \mapsto a + \alpha\, u$ is a geodesic. Then by the previous, $\Gamma^g_{a+\alpha\, u}(u,u) = 0$ for all $\alpha \in]-1,1[$. In particular for any $u_1 \in U$, by continuity of Γ^g and the fact that $U_1 \ni u = \lim_{\alpha\to 1} \alpha\, u$, $\Gamma^g_{a+u}(u,u) = \lim_{\alpha\to 1} \Gamma^g_{a+\alpha\, u}(u,u) = \lim_{\alpha\to 1} 0 = 0$, so g is simple at a.

Suppose A. We follow [Dui2006]. Let $u \in U_1$, then because scalar multiplication $\mathbb{K} \times A \to A$ and U_1 is balanced, there exists a $\delta \in]0,\infty[$ such that for all $\alpha \in B_{\mathbb{K}}(0, 1+\delta)$ we have $\alpha\, u \in U_1$. In particular $\gamma_u :]-1-\delta, 1+\delta[\to U$, $\gamma_u(\alpha) := a + \alpha\, u$ is a geodesic with respect to g with image in U_1. Hence, as $[0,1] \subseteq]-1-\delta, 1+\delta[$,

$$\begin{aligned}
\frac{1}{2}\, g_a(u,u) &= \int_0^1 \Big(\alpha \mapsto \frac{1}{2}\, g_a(u,u)\Big) \\
&= \int_0^1 \Big(\alpha \mapsto \frac{1}{2}\, g_{\gamma_u(0)}(\gamma_u'(0), \gamma_u'(0))\Big) \\
&\overset{(6.23)}{=} \int_0^1 \Big(\alpha \mapsto \frac{1}{2}\, g_{\gamma_u(\alpha)}(\gamma_u'(\alpha), \gamma_u'(\alpha))\Big) \\
&= \int_0^1 \Big(\alpha \mapsto G(\gamma_u(\alpha), \gamma_u'(\alpha))\Big),
\end{aligned}$$

so via Theorem (5.3.14) we find that for $v \in A$

$$\begin{aligned}
g_a(u,v) &= D_u\Big(u' \mapsto \frac{1}{2}\, g_a(u',u')\Big)(v) \\
&= D_u\Big(u' \mapsto \int_0^1 \Big(\alpha \mapsto G(\gamma_{u'}(\alpha), \gamma_{u'}'(\alpha))\Big)\Big)(v) \\
&\overset{(5.7)}{=} -\int_0^1 \Big(\alpha \mapsto \mathrm{La}(G,\gamma_u)\Big(\alpha, \frac{\partial \gamma_u(\alpha)}{\partial u}(v)\Big)\Big) \\
&\quad + g_{\gamma_u(1)}\Big(\gamma_u'(1), \frac{\partial \gamma_u(1)}{\partial u}(v)\Big) - g_{\gamma_u(0)}\Big(\gamma_u'(0), \frac{\partial \gamma_u(0)}{\partial u}(v)\Big) \\
&\overset{(6.22)}{=} -0 + g_{a+u}(u, 1\, v) - g_a(u, 0\, v) \\
&= g_{a+u}(u,v).
\end{aligned}$$

In particular, for all $u \in U_1$ and $v \in A$ we have

$$g_a(u,v) = g_{a+u}(u,v).$$

Taking the derivative twice with respect to u of Equation (6.24) in directions $v, w \in A$ and evaluating the result at $u = 0$, we find with Equation (5.4) that

$$0 + 0 + 0 + \Gamma^g_{a+0}(v,w) = 0,$$

so $\Gamma^g_a(v,w) = 0$ for all $v, w \in A$. Therefore, by Equation (6.5) and Equation (6.7), we find that $D_a g(w)(u,v) = 0$ for all $u, v, w \in A$. If g is simple at a, $D_a g = 0$.

Let B $/\mathbb{K}$, $V \subseteq B$ open and $\phi : V \to U$ a C^2 diffeomorphism satisfying $\phi(b) = a$ for some $b \in V$ and $D^2_{b_1}\phi = 0$ for all $b_1 \in V$. Shrink V to an open abc neighbourhood of b.

Then for any $u \in B$ the map $V \to A : b_1 \mapsto D_{b_1}\phi(u)$ is C^1 with derivative 0. Hence by Theorem (5.1.8), it is constant, so $D_{b_1}\phi(u) = D_b\phi(u)$ for all $b_1 \in V$, $u \in B$.

Let $\chi := D_b\phi$. Then by Lemma (5.1.15), $\chi : B \to A$ bijective with $\chi^{-1} : A \to B$. Now by Theorem (5.1.8), as $D_{b_1}\phi = \chi$ for all $b_1 \in V$, $\phi(b_1) = \chi(b_1 - b) + a$ for all $b_1 \in V$. Hence Equation (6.9) becomes, for all $b_1 \in \phi^{-1}(U) \cap V$ and $u, v \in B$

$$D^2_{b_1}\phi(u,v) + \Gamma^g_{\phi(b_1)}(D_{b_1}\phi(u), D_{b_1}\phi(v)) = D_{b_1}\phi(\Gamma^{\phi^* g}_{b_1}(u,v)),$$

which reduces to

$$0 + \Gamma^g_{\chi(b_1-b)+a}(\chi(u), \chi(v)) = \chi(\Gamma^{\phi^* g}_{b_1}(u,v))$$

and therefore

$$\Gamma^{\phi^* g}_{b_1}(u,v) = \chi^{-1}(\Gamma^g_{\chi(b_1-b)+a}(\chi(u), \chi(v))).$$

In particular, for all $u \in \chi^{-1}(U_1)$, $b_1 = b + u \in \phi^{-1}(U)$, so

$$\begin{aligned}\Gamma^{\phi^* g}_{b+u}(u,u) &= \chi^{-1}(\Gamma^g_{a+\chi(u)}(\chi(u), \chi(u))) \\ &= 0\end{aligned}$$

because g is simple at a and $\chi(u) \in U_1$. Hence $\phi^* g$ is simple at b.

With this, Lemma (6.3.3), and Lemma (6.3.5), we arrive at the following lemma.

Lemma 6.4.4

Let A /$\mathbb{K}$, $U \subseteq A$ open. Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A, $g \in C^2(U)$, and $a \in U$.

Then g is a simple metric at a if and only if there exists an open abc neighbourhood U_1 of 0 in $U - a$ such that for all $u \in U_1$ we have that

$$]-1,1[\to U_1 : \alpha \mapsto a + \alpha\, u$$

is a geodesic with respect to g.

If this is the case, then for all $u \in U_1$ and $v \in A$ we have [14]

$$g_a(u,v) = g_{a+u}(u,v), \tag{6.25}$$

and at a we have

$$D_a g = 0. \tag{6.26}$$

Furthermore, let B /$\mathbb{K}$, $V \subseteq B$ open. Then for any C^2 diffeomorphism $\phi : V \to U$ satisfying for all $b \in V$ that $D^2_b\phi = 0$, the pullback $\phi^* g$ is a C^2 metric that is simple at $\phi^{-1}(a)$, if g is a simple metric at a.

Note that Equation (6.26) implies that if a metric g is simple at all $a \in U$, then by Theorem (5.1.8) it must necessarily be constant on each connected component of U.

[14] This is often called Gauß's Lemma.

Suppose $g \in C^k(U)$ for $k \geq 2$. We are now going to investigate the consequences of Equation (6.25) by taking derivatives of this equation in directions $v_2, v_3, \ldots \in A$ using Equation (5.4) and Theorem (5.1.16):

$$\begin{aligned}
g_a(u, v_1) &= g_{a+u}(u, v_1) \\
g_a(v_2, v_1) &= D_{a+u}g(v_2)(u, v_1) + g_{a+u}(v_2, v_1) \\
0 &= D^2_{a+u}g(v_2, v_3)(u, v_1) + D_{a+u}g(v_2)(v_3, v_1) + D_{a+u}g(v_3)(v_2, v_1) \\
0 &= D^3_{a+u}g(v_2, v_3, v_4)(u, v_1) \\
&\quad + D^2_{a+u}g(v_2, v_3)(v_4, v_1) + D^2_{a+u}g(v_2, v_4)(v_3, v_1) + D^2_{a+u}g(v_3, v_4)(v_2, v_1) \\
&\ldots \\
0 &= D^k_{a+u}g(v_2, \ldots, v_{k+1})(u, v_1) + D^{k-1}_{a+u}g(v_2, v_3, \ldots, v_k)(v_{k+1}, v_1) \\
&\quad + D^{k-1}_{a+u}g(v_{k+1}, v_2, \ldots, v_{k-1})(v_k, v_1) + \ldots + D^{k-1}_{a+u}g(v_3, v_4, \ldots, v_{k+1})(v_2, v_1).
\end{aligned}$$

Evaluating these expressions at $u = 0$ we find for all $2 \leq l \leq k-1$ and $v_1, \ldots, v_{l+2} \in A$ that

$$D^l_a g(v_2, v_3, \ldots, v_{l+1})(v_{l+2}, v_1) + D^l_a g(v_{l+2}, v_2, \ldots, v_l)(v_{l+1}, v_1) + \ldots + D^l_a g(v_3, v_4, \ldots, v_{l+2})(v_2, v_1) = 0. \quad (6.27)$$

What we want is to express $D^l_a g(v_{l+2}, \ldots, v_3)(v_2, v_1)$ completely in terms of $\nabla^{l-2}R^g_a$ with appropriate combinations of $v_1, \ldots, v_{l+2}$ inserted in each term. [15]

To do so, it is convenient to simplify our notation using the symmetries of $D^l_a g$: by Theorem (5.1.16) and the fact that g is a metric we know that $D^l_a g(v_{l+2}, \ldots, v_3)(v_2, v_1)$ is symmetric under permutations of $v_{l+2}, \ldots, v_3$ and v_2, v_1.

Therefore, for any $\pi \in S^{l+2}$, the expression

$$D^l_a g(v_{\pi(l+2)}, \ldots, v_{\pi(3)})(v_{\pi(2)}, v_{\pi(1)})$$

is completely characterised by $\pi(1), \pi(2) \in \{1, \ldots, l+2\}$: the expression remains the same for all permutations of the indices $\pi(3), \ldots, \pi(l+2)$ by Theorem (5.1.16). Hence we will for $2 \leq l \leq k-1$, $i_1, i_2 \in \{1, \ldots, l+2\}$, $i_1 \neq i_2$ use the notation

$$(i_1, i_2)^l := D^l_a g(v_{j_1}, \ldots, v_{j_l})(v_{i_1}, v_{i_2})$$

where $j_1 < \ldots < j_l$ are values in $\{1, \ldots, l+2\} \setminus \{i_1, i_2\}$, that are uniquely determined (since there are $l + 2 - 2 = l$ elements in $\{1, \ldots, l+2\} \setminus \{i_1, i_2\}$). Then for any $\pi \in S^{l+2}$,

$$D^l_a g(v_{\pi(l+2)}, \ldots, v_{\pi(3)})(v_{\pi(2)}, v_{\pi(1)}) = (\pi(2), \pi(1))^l.$$

Note that as g is symmetric,

$$(i_1, i_2)^l = (i_2, i_1)^l,$$

and that Equation (6.27) implies

$$(1, 2)^l + (1, 3)^l + \ldots + (1, l+2)^l = 0,$$

[15] $\nabla^{l-2}R^g_a$ because R^g_a is expressed in $D^2_a g$ as highest order derivative of g by Equation (6.11) and hence, by Equation (6.15), $\nabla^{l-2}R^g_a$ contains $D^l_a g$ as highest order derivative.

which in turn gives us, by permuting the vectors $v.$, that for any $i \in \{1, \dots, l+2\}$

$$\sum_{j=1, j\neq i}^{l+2} (i,j)^l = 0. \tag{6.28}$$

Now by Equation (6.11) and Equation (6.26) we have that for any $i_1, i_2, i_3, i_4 \in \{1,2,3,4\}$

$$R_a^g(v_{i_1}, v_{i_2}, v_{i_3}, v_{i_4}) = \frac{1}{2}\Big((i_2,i_3)^2 - (i_1,i_3)^2 - (i_2,i_4)^2 + (i_1,i_4)^2\Big) + 0.$$

Hence by Equation (6.15), for $2 \leq l \leq k-1$, $j_1, \dots, j_{l-2}, i_1, \dots, i_4 \in \{1, \dots, l+2\}$

$$\nabla^{l-2} R_a^g(v_{j_1}, \dots, v_{j_{l-2}}, v_{i_1}, \dots, v_{i_4}) = \frac{1}{2}\Big((i_1,i_4)^l + (i_2,i_3)^l - (i_1,i_3)^l - (i_2,i_4)^l\Big) + \dots$$

where $\dots$ consists of lower (that is $< l$) order derivatives of g.

This can be shown using induction. For $l = 2$ we have with Equation (6.11) that for $R_a^g(v_1, \dots, v_4)$, '$\dots$' $= g_a(\Gamma_a^g(v_1,v_4), \Gamma^g(v_2,v_3)) - g_a(\Gamma_a^g(v_1,v_3), \Gamma_a^g(v_2,v_4))$ which by Equation (6.7) can be expressed entirely in g_a and $D_a g$.

Now suppose that for some $l \geq 3$ we have that for $\nabla^{(l-1)-2} R_a^g(v_1, \dots, v_{(l-1)+2})$, that '$\dots$' can be expressed in $D_a^m g$ for $0 \leq m < l-1$ and that the four remaining terms are a linear combination of $D_a^{l-1} g$ (note that this is the case for $l = 3$). We obtain $\nabla^{l-2} R_a^g(v_1, \dots, v_{l+2})$ by taking $\nabla(\nabla^{(l-1)-2} R^g)_a$, which by Equation (6.15) involves $D_a(\nabla^{(l-1)-2} R^g)$ and substituting Γ_a^g in $\nabla^{(l-1)-2} R^g$.

By substituting Γ_a^g we obtain terms that can be expressed in terms of $D_a^m g$ for $0 \leq m \leq l-1$ (Γ_a^g is determined by g_a and $D_a g$, and $\nabla^{(l-1)-2} R^g$ involves derivatives of g of at most order $l-1$ by our induction hypothesis).

$D_a(\nabla^{(l-1)-2} R^g)$ gives us the derivative of the first four non-'$\dots$' terms of $\nabla^{(l-1)-2} R^g$, which is a linear combination of four $D_a^l g$ terms, and the derivative of '$\dots$' from $\nabla^{(l-1)-2} R^g$. Since by induction, '$\dots$' from $\nabla^{(l-1)-2} R^g$ is completely expressed in terms of $D_a^m g$ for $0 \leq m < l-1$ we find that the derivative is completely expressed in terms of $D_a^m g$ for $0 \leq m \leq l-1$, where we use Equation (6.1) to write derivatives of Γ^g in terms of derivatives of g. Hence the '$\dots$' part of $\nabla^{l-2} R_a^g(v_1, \dots, v_{l+2})$ only contains terms $D_a^m g$ for $0 \leq m < l$. Now using induction we see that for $\nabla^{l-2} R_a^g$, '$\dots$' is expressed entirely in terms of $D_a^m g$ for $0 \leq m < l$ for all $2 \leq l \leq k-1$. [16]

Inspired by this and Equation (6.28) [17] we consider for $2 \leq l \leq k-1$,

[16] Note that even though $\Gamma_a^g = 0$ by Equation (6.26), the derivatives of Γ_a^g need not be 0 and repeated applications of Equation (5.4) and Theorem (5.1.8) yield very complicated expressions for all terms contained in '$\dots$'. However, all derivatives of g that occur in '$\dots$' are of lower order than l.

[17] Actually after painstakingly calculating $D_a^2(\phi^* g)(w,x)(u,v) = -(R_a^g(x,u,w,v) + R_a^g(w,u,x,v))/3$ using Theorem (6.5.2), where ϕ is the map from Theorem (6.5.1).

$v_1, \ldots, v_{l+4} \in A$ the following expression:

$$\begin{aligned}
&\sum_{\pi \in S^l} \nabla^{l-2} R^g_a(v_{\pi(1)}, \ldots, v_{\pi(l-2)}, v_{\pi(l-1)}, v_{l+1}, v_{\pi(l)}, v_{l+2}) \\
&= \frac{1}{2} \sum_{\pi \in S^l} \Big[(\pi(l-1), l+2)^l + (l+1, \pi(l))^l - (\pi(l-1), \pi(l))^l - (l+1, l+2)^l \Big] + \ldots \\
&= \frac{1}{2} \sum_{i=1}^{l} (l-1)!\, (i, l+2)^l + \frac{1}{2} \sum_{i=1}^{l} (l-1)!\, (l+1, i)^l \\
&\quad - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1, j \neq i}^{l} (l-2)!\, (i, j)^l \\
&\quad - \frac{1}{2}\, l!\, (l+1, l+2)^l + \ldots \\
&\overset{(6.28)}{=} -\frac{1}{2} (l-1)!\, (l+1, l+2)^l - \frac{1}{2} (l-1)!\, (l+1, l+2)^l \\
&\quad - \frac{1}{2} (l-2)! \sum_{i=1}^{l} \Big(-(i, l+1)^l - (i, l+2)^l \Big) \\
&\quad - \frac{1}{2}\, l!\, (l+1, l+2)^l + \ldots \\
&= -\Big(\frac{1}{2} l! + (l-1)! + (l-2)! \Big) (l+1, l+2)^l + \ldots \\
&= -\frac{l(l+1)}{2} (l-2)!\, D^l_a g(v_1, \ldots, v_l)(v_{l+1}, v_{l+2}) + \ldots
\end{aligned}$$

where again $\ldots$ denotes terms containing derivatives of g of order $< l$.

This leads us to the following lemma.

Lemma 6.4.5

Let A Vs $/\mathbb{K}$ T2 LC UC, $U \subseteq A$ open. Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A, $g \in C^k(U)$, $k \geq 2$, and $a \in U$.

Suppose g is a simple metric at a. Then for all $2 \leq l \leq k-1$ and $u, v, w_1, \ldots, w_l \in A$ we have

$$\begin{aligned}
&D^l_a g(w_1, \ldots, w_l)(u, v) = \\
&\qquad -2 \Big(\frac{l-1}{l+1} \Big) \frac{1}{l!} \sum_{\pi \in S^l} \nabla^{l-2} R^g_a(w_{\pi(1)}, \ldots, w_{\pi(l-2)}, w_{\pi(l-1)}, u, w_{\pi(l)}, v) \\
&\qquad\qquad + \ \ldots \text{expression in } D^m_a g \text{ for } 0 \leq m < l \ldots. \qquad (6.29)
\end{aligned}$$

Equation (6.29) becomes for $l = 2$ (use Equation (6.26))

$$D^2_a g(w_1, w_2)(u, v) = -\frac{1}{3} \Big(R^g_a(w_1, u, w_2, v) + R^g_a(w_2, u, w_1, v) \Big) + 0.$$

We therefore see by using induction and Equation (6.29), that the sequence $(g_a, R^g_a, \nabla R^g_a, \ldots, \nabla^{k-3} R^g_a)$ completely determines $(g_a, D_a g, D^2_a g, \ldots, D^{k-1}_a g)$ if $g \in C^k(U)$ is simple at a.

This, together with Equation (6.11) and Equation (6.15), yields the following theorem.

Theorem 6.4.6: Taylor sequences of simple metrics
Let A Vs $/\mathbb{K}$ T2 LC UC, $U \subseteq A$ open.

Let $g, h : U \times A \times A \to \mathbb{K}$ be metrics on A. Suppose $g, h \in C^k(U)$ for $k \geq 2$ and that g and h are both simple metrics at $a \in U$.

Then $D_a^l g = D_a^l h$ for all $0 \leq l < k$ if and only if $g_a = h_a$ and $\nabla^l R_a^g = \nabla^l R_a^h$ for all $0 \leq l < k - 2$.

A similar result, obtained in an entirely different fashion can be found in [Car1951], Chapitre X, n° 218 - 219 on page 238.

6.5 Making metrics simple

Now that we know, in the form of Theorem (6.4.6), that the Taylor sequence of a metric g which is simple at a is completely determined by g_a and $\nabla^l R_a^g$, we need to ascertain whether or not a general metric can be brought into a simple form. [18]

Theorem 6.5.1
Let A Ba $/\mathbb{K}$, $U \subseteq A$ open.

Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A and suppose that $g \in C^k(U)$ for $k \geq 3$.

Then for any $a \in U$ there exists an open neighbourhood U_1 of a in U, an open abc neighbourhood V of 0 in A, and a C^{k-1} diffeomorphism $\phi : V \to U_1$ such that $\phi(0) = a$, $D_0\phi = \mathrm{id}_A$, and $\phi^* g$ is a simple metric at 0.

Proof. Using the diffeomorphism $A \to A : a_1 \mapsto a_1 - a$ we see that we may take $a = 0$ without loss of generality.

Define $f : U \times A \to A \times A$ by

$$f(a, u) := (u, -\Gamma_a^g(u, u)). \tag{6.30}$$

Since $g \in C^k(U)$, $\Gamma^g \in C^{k-1}(U \times A \times A, A)$, so $f \in C^{k-1}(U \times A, A \times A)$. Furthermore, as $k \geq 3$, f is C^2 and hence by Theorem (5.5.10) the flow $e^{\cdot\, f}$ exists on a certain open set $W \subseteq \mathbb{R} \times U \times A$ with $\{0\} \times U \times A \subseteq W$.

For convenience we define the maps $\pi_1, \pi_2 : A \times A \to A$, $\iota : A \to A \times A$ by

$$\pi_1(a_1, a_2) := a_1, \qquad \pi_2(a_1, a_2) := a_2, \qquad \iota(a) := (0, a). \tag{6.31}$$

Note that π_1, π_2, and ι are C 1 and hence (Theorem (5.1.8)) C^∞. We define for $S(0, u) \neq \emptyset$ the curve $\gamma_u : S(0, u) \to U$ to be

$$\gamma_u(\alpha) := \pi_1(e^{\alpha\, f}(0, u)).$$

Equation (5.10) and our expression for f then give us that γ_u is the unique maximal solution to [19]

$$\gamma_u(0) = 0, \qquad \gamma_u'(0) = u, \qquad \gamma_u''(\alpha) + \Gamma^g_{\gamma_u(\alpha)}(\gamma_u'(\alpha), \gamma_u'(\alpha)) = 0, \tag{6.32}$$

[18] The answer is yes, see Theorem (6.5.1), by pulling back the metric with the exponential map (see [Ban2008], Section 9, page 33) induced by the geodesics. This yields what is referred to as the metric in 'Riemann normal coordinates'.

[19] As for any open interval $S \subseteq \mathbb{R}$ and C^1 map $S \to U \times A : \alpha \mapsto (\gamma(\alpha), \eta(\alpha))$ we have $(\gamma, \eta)'(\alpha) = f((\gamma, \eta)(\alpha))$ if and only if $(\gamma'(\alpha), \eta'(\alpha)) = (\eta(\alpha), -\Gamma_{\gamma(\alpha)}(\eta(\alpha), \eta(\alpha)))$ if and only if $\gamma'(\alpha) = \eta(\alpha)$ and $\gamma''(\alpha) = \eta'(\alpha) = -\Gamma_{\gamma(\alpha)}(\gamma'(\alpha), \gamma'(\alpha))$.

in particular γ_u is a geodesic. Note that with this definition

$$e^{\alpha\, f}(0,u) = (e^{\alpha\, f} \circ \iota)(u) = (\gamma_u(\alpha), \gamma_u'(\alpha))$$

and γ_u is C^{k-1}.

Define for $u \in A$ with $S(0,u) \neq \emptyset$ and $\beta \in \mathbb{K}$, $\beta \neq 0$ the map $\eta : \frac{1}{\beta}\, S(0,u) \to U$ by $\eta(\alpha) := \gamma_u(\beta\,\alpha)$. Then $\eta'(\alpha) = \beta\,\gamma_u'(\beta\,\alpha)$ and $\eta''(\alpha) = \beta^2\,\gamma_u''(\beta\,\alpha)$, so $\eta(0) = 0$, $\eta'(0) = \beta\,u$ and

$$\begin{aligned}\eta''(\alpha) + \Gamma^g_{\eta(\alpha)}(\eta'(\alpha),\eta'(\alpha)) &= \beta^2\,\gamma_u''(\beta\,\alpha) + \Gamma^g_{\gamma_u(\beta\,\gamma)}(\beta\,\gamma_u'(\beta\,\alpha), \beta\,\gamma_u'(\beta\,\alpha))\\ &= \beta^2\, 0\\ &= 0.\end{aligned}$$

Hence (by Theorem (5.5.10), $\gamma_{\beta\,u}$ is the unique maximal solution) $\eta(\alpha) = \gamma_{\beta\,u}(\alpha)$ and $\frac{1}{\beta}\, S(0,u) \subseteq S(0,\beta\,u)$. Similarly we find $\beta\, S(0,\beta\,u) \subseteq S(0,u)$.

So for all $u \in A$ with $S(0,u) \neq \emptyset$, $\beta \in \mathbb{K}$, $\beta \neq 0$, and $\alpha \in \frac{1}{\beta}\, S(0,u)$ we have

$$S(0,u) = \beta\, S(0,\beta\,u), \qquad \gamma_u(\beta\,\alpha) = \gamma_{\beta\,u}(\alpha). \tag{6.33}$$

Now consider the set

$$V := \{u \in A \mid (1,0,u) \in W\} \subseteq A.$$

Then first of all, $0 \in W$ as $\gamma_0(\alpha) = 0$ for all $\alpha \in \mathbb{R}$ satisfies Equation (6.32) and V is open because W is open. So V is an open neighbourhood of 0 in A.

This permits us to define the map $\phi : V \to U$ by

$$\phi(u) := \pi_1(e^{1\, f}(0,u)) = (\pi_1 \circ e^{1\, f} \circ \iota)(u).$$

Since f is C^{k-1}, $e^{\cdot\, f}$ is C^{k-1} by Theorem (5.5.10), π_1 is C^∞ as continuous linear map, so ϕ is C^{k-1}. Furthermore, for all $u \in V$, $\alpha \in \mathbb{K}$ with $\alpha\,u \in V$ we have

$$\phi(\alpha\,u) = \gamma_{\alpha\,u}(1) = \gamma_u(\alpha)$$

by Equation (6.33).

As $\phi(0) = 0$ and ϕ is C^2, for any $u \in A$ by Lemma (5.1.4)

$$\begin{aligned}D_0\phi(u) &= \lim_{\alpha\to 0} \frac{1}{\alpha}\Big(\phi(\alpha\,u) - \phi(0)\Big)\\ &= \lim_{\alpha\to 0} \frac{1}{\alpha}\Big(\gamma_u(\alpha) - \gamma_u(0)\Big)\\ &= \gamma_u'(0)\\ &= u.\end{aligned}$$

So ϕ is C^2 and $D_0\phi = \mathrm{id}_A$ which is invertible, therefore by Theorem (5.5.8) we can shrink V to an open abc neighbourhood of 0 in A and find an open neighbourhood U_1 of 0 in U such that $\phi : V \to U_1$ is a C^{k-1} diffeomorphism.

Let $u \in A$ and $\alpha \in \mathbb{K}$ such that $\alpha\,u \in V$. Then

$$\begin{aligned}D_{\alpha\,u}\phi(u) &= \lim_{\beta\to 0} \frac{1}{\beta}\Big(\phi(\alpha\,u + \beta\,u) - \phi(\alpha\,u)\Big)\\ &= \lim_{\beta\to 0} \frac{1}{\beta}\Big(\gamma_u(\alpha+\beta) - \gamma_u(\alpha)\Big)\\ &= \gamma_u'(\alpha),\end{aligned}$$

and likewise

$$\begin{aligned}
D^2_{\alpha\, u}\phi(u,u) &= \lim_{\beta\to 0}\frac{1}{\beta}\Big(D_{\alpha\, u+\beta\, u}\phi(u) - D_{\alpha\, u}\phi(u)\Big)\\
&= \lim_{\beta\to 0}\frac{1}{\beta}\Big(\gamma_u'(\alpha+\beta) - \gamma_u'(\alpha)\Big)\\
&= \gamma_u''(\alpha).
\end{aligned}$$

Note that as ϕ is C^2 and g is C^1, $\phi^* g$ is a C^1 metric on V by Lemma (6.3.3) and Lemma (6.3.5).

By Equation (6.9) we furthermore obtain for all $u \in V$ that

$$\begin{aligned}
D_u\phi(\Gamma_u^{\phi^* g}(u,u)) &= D^2_u\phi(u,u) + \Gamma^g_{\phi(u)}(D_u\phi(u), D_u\phi(u))\\
&= \gamma_u''(1) + \Gamma^g_{\gamma_u(1)}(\gamma_u'(1),\gamma_u'(1))\\
&\overset{(6.32)}{=} 0.
\end{aligned}$$

As $D_u\phi$ is bijective (because $\phi|_V$ is a diffeomorphism), we therefore find that for all $u \in V$, $\Gamma_u^{\phi^* g}(u,u) = 0$: $\phi^* g$ is a simple metric at 0. $\square$

We will now give a means by which we can determine ϕ's Taylor series, using Lemma (5.5.12).

Let $u \in V$, $0 \le l < k-1$, $v_1,\dots,v_l \in A$. Note that as $\phi(u) = \pi_1(e^{1\,f}(0,u))$, by Theorem (5.1.8) we have (π_1, ι are C I)

$$\begin{aligned}
D^l_u\phi(v_1,\dots,v_l) &= D^l_{(0,u)}(\pi_1\circ e^{1\,f})((0,v_1),\dots,(0,v_l))\\
&= \pi_1(D^l_u(e^{1\,f}\circ\iota)(v_1,\dots,v_l)).
\end{aligned}$$

This and Lemma (5.5.12) motivate us to choose $\eta : S(0,u) \to A\times A$ as

$$\eta(\alpha) := (\eta_1(\alpha),\eta_2(\alpha)) := D^l_u(e^{\alpha\, f}\circ\iota)(v_1,\dots,v_l),$$

such that

$$D^l_u\phi(v_1,\dots,v_l) = \eta_1(1).$$

As $\eta(\alpha) = D^l_{(0,u)}e^{\alpha\, f}((0,v_1),\dots,(0,v_l))$, Lemma (5.5.12) gives us

$$\begin{aligned}
\eta(0) &= \begin{cases} (0,u) & l=0\\ (0,v_1) & l=1\\ (0,0) & l>1\end{cases}\\
\eta'(\alpha) &= D^l_{(0,u)}(f\circ e^{\alpha\, f})((0,v_1),\dots,(0,v_l)).
\end{aligned}$$

With Equation (6.30) we find that for the first term $\pi_1(f(a,u)) = u$, all derivatives of higher than first order vanish and (π_1 is C I) $\pi_1(D_{(a,u)}f(v,w)) = D_{(a,u)}(\pi_1\circ f)(v,w) = w$. Hence, if we work out $\eta_1'(\alpha)$ using Theorem (5.1.8) and Equation (5.4), we find for $l > 0$ that

$$
\begin{aligned}
\eta_1'(\alpha) &= \pi_1\Big(D^l_{(0,u)}(f \circ e^{\alpha f})((0,v_1),\dots,(0,v_l))\Big) \\
&= D^{l-1}_{(0,u)}((a',u') \mapsto D_{e^{\alpha f}(a',u')}(\pi_1 \circ f)(D_{(a',u')}e^{\alpha f}(0,v_l)))((0,v_1),\dots,(0,v_{l-1})) \\
&= \dots \text{ (use Equation (5.4) repeatedly)} \\
&= D_{e^{\alpha f}(0,u)}(\pi_1 \circ f)(D^l_{(0,u)}e^{\alpha f}((0,v_1),\dots,(0,v_l))) \\
&\quad + \text{ (derivatives of } \pi_1 \circ f \text{ of order at least 2)} \\
&= D_{e^{\alpha f}(0,u)}(\pi_1 \circ f)(\eta(\alpha)) \\
&\quad + 0 \\
&= \eta_2(\alpha),
\end{aligned}
$$

and also for $l = 0$

$$
\begin{aligned}
\eta_1'(\alpha) &= \pi_1(f(e^{\alpha f}(0,u))) \\
&= \pi_1(f(\eta(\alpha))) \\
&= \eta_2(\alpha).
\end{aligned}
$$

So $\eta_1'(\alpha) = \eta_2(\alpha)$, therefore we find that

$$
\eta_1(0) = 0, \qquad \eta_1'(0) = \eta_2(0) = \begin{cases} u & l = 0 \\ v_1 & l = 1 \\ 0 & l > 1 \end{cases},
$$

as well as

$$
\begin{aligned}
\eta_1''(\alpha) &= \eta_2'(\alpha) \\
&= \pi_2(D^l_{(0,u)}(f \circ e^{\alpha f})((0,v_1),\dots,(0,v_l))) \\
&= D^l_u(\pi_2 \circ f \circ e^{\alpha f} \circ \iota)(v_1,\dots,v_l).
\end{aligned}
$$

Note that for $l = 0$ we precisely obtain Equation (6.21) for $\gamma_u = \eta_1$. This leads us to the following theorem.

Theorem 6.5.2: Derivatives of ϕ
Let A Ba $/\mathbb{K}$, $U \subseteq A$ open.

Let $g : U \times A \times A \to \mathbb{K}$ be a metric on A and suppose that $g \in C^k(U)$ for $k \geq 3$.

Then the map ϕ from Theorem (6.5.1) has the following property.

For all $0 \leq l < k-1$, $u \in V$, $v_1, \dots, v_l \in A$ the curve $\gamma_{u,v_1,\dots,v_l} : S(0,u) \to U$ defined by

$$
\gamma_{u,v_1,\dots,v_l}(\alpha) := \pi_1(D^l_u(e^{\alpha f} \circ \iota)(v_1,\dots,v_l))
$$

(where f is given by Equation (6.30) and π_1, π_2, and ι by Equation (6.31)) satisfies

$$
\begin{aligned}
\gamma_{u,v_1,\dots,v_l}(0) &= 0 \\
\gamma'_{u,v_1,\dots,v_l}(0) &= \begin{cases} u & l = 0 \\ v_1 & l = 1 \\ 0 & l > 1 \end{cases}
\end{aligned}
$$

and for all $\alpha \in S(0,u)$

$$\gamma''_{u,v_1,\ldots,v_l}(\alpha) = D^l_u(\pi_2 \circ f \circ e^{\alpha\, f} \circ \iota)(v_1,\ldots,v_l). \tag{6.34}$$

These curves completely determine ϕ by

$$D^l_u\phi(v_1,\ldots,v_l) = \gamma_{u,v_1,\ldots,v_l}(1).$$

Note that for $l = 0$ we obtain geodesics γ_u as solutions, and for $l > 0$, $(v_1,\ldots,v_l) \mapsto \gamma_{u,v_1,\ldots,v_l}$ is l-1 by linearity of $\pi_1 \circ D^l_u(e^{\alpha\, f} \circ \iota)$.

Noting that $\gamma_0(\alpha) = 0$ for all $\alpha \in \mathbb{R}$ is a solution, we see that Equation (6.34) simplifies considerably for $u = 0$. Using this, it is straightforward to determine that $\gamma_{0,v_1}(\alpha) = \alpha\, v_1$ (which gives $D_0\phi(v_1) = v_1$), $\gamma_{0,v_1,v_2}(\alpha) = -\alpha^2\, \Gamma^g_0(v_1,v_2)$ (which gives $D^2_0\phi(v_1,v_2) = -\Gamma^g_0(v_1,v_2)$). After a slightly longer calculation, we find

$$\begin{aligned}
\gamma_{0,v_1,v_2,v_3}(\alpha) = \frac{1}{3}\alpha^3\Big(&2\,\Gamma^g_0(v_1,\Gamma^g_0(v_2,v_3)) + 2\,\Gamma^g_0(v_2,\Gamma^g_0(v_3,v_1))\\
&+ 2\,\Gamma^g_0(v_3,\Gamma^g_0(v_1,v_2)) - D_0\Gamma^g(v_1)(v_2,v_3)\\
&- D_0\Gamma^g(v_2)(v_3,v_1) - D_0\Gamma^g(v_3)(v_1,v_2)\Big),\\
D^3_0\phi(v_1,v_2,v_3) = \frac{1}{3}\Big(&2\,\Gamma^g_0(v_1,\Gamma^g_0(v_2,v_3)) + 2\,\Gamma^g_0(v_2,\Gamma^g_0(v_3,v_1))\\
&+ 2\,\Gamma^g_0(v_3,\Gamma^g_0(v_1,v_2)) - D_0\Gamma^g(v_1)(v_2,v_3)\\
&- D_0\Gamma^g(v_2)(v_3,v_1) - D_0\Gamma^g(v_3)(v_1,v_2)\Big),
\end{aligned}$$

and in principle $D^l_0\phi(v_1,\ldots,v_l)$ may be calculated in this fashion up to any desired order $l < k$. [20]

6.6 Digression (cont'd)

We can now combine Theorem (6.4.6) and Theorem (6.5.1) into one theorem.

Theorem 6.6.1
Let A, B /$\mathbb{K}$, $U \subseteq A$, $V \subseteq B$ open. Let $k \in \mathbb{N}$, and $g : U \times A \times A \to \mathbb{K}$, $g \in C^{k+2}(U)$, $h : V \times B \times B \to \mathbb{K}$, $h \in C^{k+2}(V)$ be metrics on A and B respectively. Let $a \in U$, $b \in V$, and suppose that U and V are C^{k+3} diffeomorphic.

Then the following two statements are equivalent.

- There exists an open abc neighbourhood W of 0 in A and C^{k+1} diffeomorphisms $\phi : W \to U_1$, $\chi : W \to V_1$, where $a \in U_1 \subseteq U$, $b \in V_1 \subseteq V$ and U_1, V_1 are open, such that for all $0 \le l < k$ we have

$$D^l_0(\phi^* g) = D^l_0(\chi^* h) \tag{6.35}$$

  and $\phi^* g$ and $\chi^* h$ are simple metrics at 0.

[20] The number of terms arising from the derivative in Equation (6.34) will rapidly increase however.

- There exists a map $f : A \to B$ C l and bijective such that for all $0 \leq l < k-2$ we have for all $u_1, \ldots, u_{l+4} \in A$ that

$$\begin{aligned} g_a(u_1, u_2) &= h_b(f(u_1), f(u_2)) \\ \nabla^l R^g_a(u_1, \ldots, u_{l+4}) &= \nabla^l R^h_b(f(u_1), \ldots, f(u_{l+4})). \end{aligned} \tag{6.36}$$

Proof. Let $a \in U$, $b \in V$, and let $\psi : U \to V$ be a C^{k+3} diffeomorphism. Then by Lemma (6.3.3) and Lemma (6.3.5), $\psi^* h : U \times A \times A \to \mathbb{K}$ is a C^{k+2} metric on A.

Suppose that for $f : A \to B$ C l and bijective, f, g, and h satisfy Equation (6.36).

As A Ba, g and $\psi^* h$ both satisfy the conditions of Theorem (6.5.1) at the points a and $\psi^{-1}(b)$ respectively. Hence there exist open neighbourhoods U_1 and U_2 of a and $\psi^{-1}(b)$ in U, open abc neighbourhoods W_1 and W_2 of 0 in A, and C^{k+2-1} diffeomorphisms $\phi : W_1 \to U_1$, $\omega : W_2 \to U_2$ such that $\phi^* g$ and $\omega^*(\psi^* h)$ are both simple metrics at 0, $\phi(0) = a$, $\omega(0) = \psi^{-1}(b)$, and $D_0\phi = D_0\omega = \mathrm{id}_A$. Let $V_1 := \psi(U_2) \subseteq \psi(U) = V$, then V_1 is open since ψ^{-1} C. Note that by Lemma (6.3.4), $\omega^*(\psi^* h) = (\psi \circ \omega)^* h$.

Now $(\psi \circ \omega)(0) = \psi(\psi^{-1}(b)) = b$ and by Theorem (5.1.8) $D_0(\psi \circ \omega) = D_{\omega(0)}\psi \circ D_0\omega = D_{\psi^{-1}(b)}\psi$. Hence by Equation (6.16)

$$((\psi \circ \omega)^* h)_0(u, v) = h_b(D_{\psi^{-1}(b)}\psi(u), D_{\psi^{-1}(b)}\psi(v))$$

for all $u, v \in A$. Note that by Theorem (5.1.8) and Theorem (4.4.3), $f : A \to B$ is a C^∞ diffeomorphism, similarly $D_{\psi^{-1}(b)}\psi : A \to B$ is a C^∞ diffeomorphism with inverse (use Lemma (5.1.15)) $D_b\psi^{-1} : B \to A$. Because we would like to use Theorem (6.4.6) we consider the map

$$\chi := \psi \circ \omega \circ D_b\psi^{-1} \circ f : (D_b\psi^{-1} \circ f)^{-1}(W_2) \to V_1,$$

which is a C^{k+1} diffeomorphism. Then $\chi(0) = (\psi \circ \omega)(0) = b$ and $D_0\chi = D_{\psi^{-1}(b)}\psi \circ D_b\psi^{-1} \circ f = f$ by Theorem (5.1.8) and Lemma (5.1.15).

Restrict both ϕ and χ to $W := W_1 \cap (D_b\psi^{-1} \circ f)^{-1}(W_2)$ which is an open abc neighbourhood of 0 in A by Lemma (4.3.5) (as f and $D_b\psi^{-1}$ are C l and bijective), and shrink U_1 and V_1 to $\phi(W)$ and $\chi(W)$ respectively. By Lemma (6.4.4), both $\phi^* g$ and $\chi^* h$ are simple metrics at 0 (since $(\psi \circ \omega)^* h$ is a simple metric at 0 and $D^2_{a_1}(D_b\psi^{-1} \circ f) = 0$ for all $a_1 \in A$).

Then by Lemma (6.3.4) and our assumption on f

$$\begin{aligned} (\phi^* g)_0(u, v) &= g_{\phi(0)}(D_0\phi(u), D_0\phi(v)) \\ &= g_a(u, v) \\ &\overset{(6.36)}{=} h_b(f(u), f(v)) \\ &= h_{\chi(0)}(D_0\chi(u), D_0\chi(v)) \\ &= (\chi^* h)_0(u, v) \end{aligned}$$

for all $u, v \in A$.

By Theorem (6.2.3) we know that for all $0 \leq l < k-2$, $u_1, \ldots, u_{l+4} \in A$ we have

$$\begin{aligned}
\nabla^l R_0^{\phi^* g}(u_1, \ldots, u_{l+4}) &= \nabla^l R^g_{\phi(0)}(D_0\phi(u_1), \ldots, D_0\phi(u_{l+4})) \\
&= \nabla^l R^g_a(u_1, \ldots, u_{l+4}) \\
&\overset{(6.36)}{=} \nabla^l R^h_b(f(u_1), \ldots, f(u_{l+4})) \\
&= \nabla^l R^h_{\chi(0)}(D_0\chi(u_1), \ldots, D_0\chi(u_{l+4})) \\
&= \nabla^l R_0^{\chi^* h}(u_1, \ldots, u_{l+4}).
\end{aligned}$$

Now we can use Theorem (6.4.6) and the fact that $\phi^* g$ and $\chi^* h$ are C^k metrics by Lemma (6.3.3) to conclude that for all $0 \leq l < k$, we have that $D_0^l(\phi^* g) = D_0^l(\chi^* h)$: the maps ϕ and χ satisfy the conditions of the first item.

Suppose conversely that two C^{k+1} diffeomorphisms ϕ and χ satisfy all conditions of the first item. Then by Theorem (6.4.6) we have that $(\phi^* g)_0(u, v) = (\chi^* h)_0(u, v)$ and for all $0 \leq l < k-2$ and $u_1, \ldots, u_{l+4} \in A$, $\nabla^l R_0^{\phi^* g}(u_1, \ldots, u_{l+4}) = \nabla^l R_0^{\chi^* h}(u_1, \ldots, u_{l+4})$. We therefore pick

$$f := D_0\chi \circ [D_0\phi]^{-1} : A \to B$$

which is c l and bijective by Lemma (5.1.15).

Then by Theorem (6.2.3) and Equation (6.16), for all $0 \leq l < k-2$, $u_1, \ldots, u_{l+4} \in A$ we have

$$\begin{aligned}
\nabla^l R^g_a(u_1, \ldots, u_{l+4}) &= \nabla^l R_0^{\phi^* g}([D_0\phi]^{-1}(u_1), \ldots, [D_0\phi]^{-1}(u_{l+4})) \\
&= \nabla^l R_0^{\chi^* h}([D_0\phi]^{-1}(u_1), \ldots, [D_0\phi]^{-1}(u_{l+4})) \\
&= \nabla^l R^h_{\chi(0)}(D_0\chi([D_0\phi]^{-1}(u_1)), \ldots, D_0\chi([D_0\phi]^{-1}(u_{l+4}))) \\
&= \nabla^l R^h_b(f(u_1), \ldots, f(u_{l+4})).
\end{aligned}$$

Similarly

$$\begin{aligned}
g_a(u, v) &= (\phi^* g)_0([D_0\phi]^{-1}(u), [D_0\phi]^{-1}(v)) \\
&= (\chi^* h)_0([D_0\phi]^{-1}(u), [D_0\phi]^{-1}(v)) \\
&= h_b(f(u), f(v)).
\end{aligned}$$

So f satisfies Equation (6.36) and the second item is satisfied. □

Theorem (6.6.1) is the partial converse to Christoffel's result that was mentioned in the introduction. It is a converse in the sense that while Christoffel showed in [Chr1869] that two metrics g and h related via a diffeomorphism[21] give a chain of relations between R^g and R^h, ∇R^g and ∇R^h, …in the form of Equation (6.14) (Theorem (6.2.3)), we obtain in Theorem (6.6.1) from a chain of relations in the form of Equation (6.14), equality of the derivatives of simple forms of g and h. Of course Theorem (6.2.3) immediately gives the relations for all points in U and V (as all these points can be related by the diffeomorphism which relates g and h), while Theorem (6.6.1) only gives the 'converse' at two selected points; we have no diffeomorphism which tells us what points a and b

[21] Like in Equation (6.2).

to compare to one another. Nevertheless, it is striking that we can still recover such an amount of information about the behaviour of g from R^g, ∇R^g, ...

This is also as far as we will get in this thesis to attaining the goal stated at the beginning of Section 6.2.

We now return to discuss the questions raised about the action of $\mathcal{G}_U^k$ on $\mathcal{M}_U^l$ for $k > l$, given by Equation (6.20), to look at Theorem (6.4.6) and Theorem (6.5.1) from a different point of view.

For convenience, let

$$\mathcal{U} := \{U \subseteq A |\ U \text{ open}, 0 \in U\ \}$$

be the collection of open neighbourhoods of 0.

It turns out to be inconvenient to restrict oneself to diffeomorphisms $U \to U$, we therefore consider the *groupoid* [22] of C^k diffeomorphisms that leave 0 fixed, defined for $k \in \mathbb{N}$ by

$$\mathcal{G}^k := \{\phi : U_1 \to U_2 |\ \phi\ C^k \text{ diffeomorphism}, \phi(0) = 0, U_1, U_2 \in \mathcal{U}\ \}.$$

Multiplication is defined as $\phi\,\chi := \phi \circ \chi$ whenever $\mathrm{im}(\chi) = \mathrm{dom}(\phi)$ and inversion by taking the inverse (as a function) of the diffeomorphism, this makes $\mathcal{G}^k$ a groupoid.

We also define

$$\mathcal{M}^k := \bigcup_{U \in \mathcal{U}} \mathcal{M}_U^k$$

as the collection of all C^k metrics defined on an open neighbourhood of 0.

Now $\mathcal{G}^k$ has the same 'action' [23] on $\mathcal{M}^k$ as defined in Equation (6.20): for $k > l$, $\phi \in \mathcal{G}^k$, $g \in \mathcal{M}^l$, $g \in C^l(U)$ we define $\phi \cdot g := \phi_* g$ whenever $\mathrm{dom}(\phi) = U$. Note that this agrees with the groupoid multiplication, if $\phi \cdot g$ is defined, then $\chi \cdot \phi$ is defined if and only if $\mathrm{im}(\phi) = \mathrm{dom}(\chi)$ if and only if $\chi \cdot (\phi \cdot g)$ is defined. Furthermore, by Lemma (6.3.4), if this is the case then $(\chi\,\phi) \cdot g = \chi \cdot (\phi \cdot g)$.

Let us now formalise the 'considering of the Taylor series' we did at the end of Section 6.3. Let A Ba $/\mathbb{K}$.

First we need a container for Taylor sequences of diffeomorphisms $\phi \in \mathcal{G}^k$ for $k \in \mathbb{N}$.

Define $\mathcal{G}_0^{k,1}$ to be the collection of all $f : A \to A$ c l $/\mathbb{K}$ that are bijective. For $l > 1$ we define $\mathcal{G}_0^{k,l}$ to be the collection of all $f : A^l \to A$ c l-l $/\mathbb{K}$ that are symmetric: $f(a_1, \ldots, a_l) = f(a_{\pi(1)}, \ldots, a_{\pi(l)})$ for all $\pi \in S^l$ and $a_1, \ldots, a_l \in A$. Then we define

$$\mathcal{G}_0^k := \bigcup_{l=1}^{k} \mathcal{G}_0^{k,l}.$$

By Theorem (5.1.16) and Lemma (5.1.15) we can for all $k \in \mathbb{N}$ create a well-defined mapping (recall that for $\phi \in \mathcal{G}^k$ we have $\phi(0) = 0$, so we discard the first term of ϕ's Taylor series)

$$\rho^k : \mathcal{G}^k \to \mathcal{G}_0^k : \phi \mapsto (D_0\phi, D_0^2\phi, \ldots, D_0^k\phi).$$

[22] A groupoid is a group in the usual sense of the word, except that the product need not be defined for all pairs of elements of the groupoid.

[23] Just as the multiplication of the groupoid, this is the same as a group action which is not defined for all elements of the groupoid and the set on which it acts.

It is clear that ρ^k is not injective for any $k \in \mathbb{N}$ (by considering diffeomorphisms with different $(k+1)$-th derivative), which is not surprising since we discard a lot of information by only looking at the Taylor series.

Let $(f_1, f_2, \ldots, f_k) \in \mathcal{G}_0^k$ be given. Define $\phi : A \to A : a \mapsto f_1(a) + f_2(a,a) + \ldots + f_k(a, \ldots, a)$, then ϕ is C^∞ (as all the f_l are C^∞ because they are ⓒ l-ⓛ), $\phi(0) = 0$, and $D_0\phi = f_1$. By definition of $\mathcal{G}_0^{k,1}$ this makes $D_0\phi$ ⓒⓛ and bijective, so by Theorem (5.5.8) there exist open neighbourhoods U_1 and U_2 of 0 and $\phi(0) = 0$ in A such that $\phi|_{U_1} : U_1 \to U_2$ is a C^∞ diffeomorphism. Hence $\phi|_{U_1} \in \mathcal{G}^k$ and because all f_l are symmetric, $\rho^k(\phi|_{U_1}) = (f_1, f_2, \ldots, f_k)$. Therefore ρ^k is surjective.

We endow $\mathcal{G}_0^k$ with a multiplication map $\mathcal{G}_0^k \times \mathcal{G}_0^k \to \mathcal{G}_0^k$ (using surjectivity of ρ^k to write all f_l as the $D_0^l\phi$ of a certain ϕ)

$$\begin{aligned}
&(D_0\chi, D_0^2\chi, \ldots)\,(D_0\phi, D_0^2\phi, \ldots) \\
&:= \Big(u_1 \mapsto D_0\chi(D_0\phi(u_1)), \\
&\qquad (u_1, u_2) \mapsto D_0^2\chi(D_0\phi(u_1), D_0\phi(u_2)) + D_0\chi(D_0^2\phi(u_1,u_2)), \ldots\Big)
\end{aligned}$$

given by repeated use of Theorem (5.1.8) and Equation (5.4) to write out $D_0^l(\chi \circ \phi)$ for $1 \leq l \leq k$. [24] This definition ensures that $\rho^k(\chi\,\phi) = \rho^k(\chi)\,\rho^k(\phi)$ whenever $\chi\,\phi$ is well-defined.

Note that $\mathcal{G}_0^1$ is precisely the collection of all bijective ⓒⓛ mappings $A \to A$, which is a group under composition and inversion of mappings. Let for $k \in \mathbb{N}$

$$\pi_1^k : \mathcal{G}_0^k \to \mathcal{G}_0^1 : (D_0\phi, D_0^2\phi, \ldots, D_0^k\phi) \mapsto D_0\phi$$

which is clearly surjective and not injective for $k > 1$. Then this map satisfies

$$\begin{aligned}
&\pi_1^k((D_0\chi, D_0^2\chi, \ldots, D_0^k\chi)\,(D_0\phi, D_0^2\phi, \ldots, D_0^k\phi)) \\
&= \pi_1^k(D_0\chi \circ D_0\phi, \ldots) \\
&= D_0\chi \circ D_0\phi \\
&= D_0\chi\, D_0\phi
\end{aligned}$$

and is therefore compatible with multiplication on $\mathcal{G}_0^k$. Note that in particular $(\pi_1^k \circ \rho^k)(\chi\,\phi) = (\pi_1^k \circ \rho^k)(\chi)\,(\pi_1^k \circ \rho^k)(\phi)$ whenever $\chi\,\phi$ is defined in $\mathcal{G}^k$.

We will now construct a similar set for $\mathcal{M}^k$, $k \in \mathbb{N}$.

Define $\mathcal{M}_0^{k,0}$ to be the collection of all $f : A^2 \to \mathbb{K}$ ⓒ 2-ⓛ /$\mathbb{K}$ that are symmetric, $f(a_1, a_2) = f(a_2, a_1)$, and satisfy that $A \to A' : a_1 \mapsto (a_2 \mapsto f(a_1, a_2))$ is bijective (recall Definition (6.1.2)). For $l > 0$ we define $\mathcal{M}_0^{k,l}$ to consist of all $f : A^{l+2} \to A$ ⓒ $(l+2)$-ⓛ /$\mathbb{K}$ that are symmetric with respect to the first l and last 2 variables: $f(a_1, \ldots, a_l, a_{l+1}, a_{l+2}) = f(a_1, \ldots, a_l, a_{l+2}, a_{l+1}) = f(a_{\pi(1)}, \ldots, a_{\pi(l)}, a_{l+1}, a_{l+2})$ for all $\pi \in S^l$ and $a_1, \ldots, a_{l+2} \in A$. Then we define

$$\mathcal{M}_0^k := \bigcup_{l=0}^{k} \mathcal{M}_0^{k,l}.$$

[24]Note that an inversion map is not easily obtained in $\mathcal{G}_0^k$, so we will not consider this set as a group, just as a set with a multiplication map $\mathcal{G}_0^k \times \mathcal{G}_0^k \to \mathcal{G}_0^k$ which is associative, because $D_0^l((\psi \circ \chi) \circ \phi) = D_0^l(\psi \circ (\chi \circ \phi)) = D_0^l(\psi \circ \chi \circ \phi)$.

Definition (6.1.2) and Theorem (5.1.16) ensure that for all $k \in \mathbb{N}$ we can create a well-defined mapping

$$\sigma^k : \mathcal{M}^k \to \mathcal{M}_0^k : g \mapsto (g_0, D_0 g, \ldots, D_0^k g).$$

As with ρ_U^k, σ_U^k is not injective.

Let $(f_0, f_1, \ldots, f_k) \in \mathcal{M}_0^k$ be given. Define $g : A \times A \times A \to \mathbb{K} : (a, u, v) \mapsto f_0(u, v) + f_1(a, u, v) + \ldots + f_k(a, \ldots, a, u, v)$. Then g is a symmetric 2-tensor, $g \in C^\infty(A)$, and $\hat{g}_0 = \Big(u \mapsto \big(v \mapsto f_0(u, v)\big)\Big)$ which is Ⓒ① $/\mathbb{K}$ and bijective $A \to A'$ by the fact that $f_0 \in \mathcal{M}_0^{k,0}$. In particular by Corollary (4.4.6), $A \simeq A'$ via $\hat{g}_0$, so A' Ba $/\mathbb{K}$. This gives us that $\hat{g}_0 \in L(A, A')^*$, so by Theorem (5.5.6) and the fact that $A \to L(A, A') : a \mapsto \hat{g}_a$ Ⓒ (as g is C^∞, use Corollary (5.5.7)) we obtain that there is an open neighbourhood U of 0 in A such that for all $a \in U$, $\hat{g}_a \in L(A, A')^*$. Then $\hat{g}^{-1}$ Ⓒ, because inversion is continuous by Theorem (5.5.6). Therefore $g|_{U \times A \times A}$ is a metric and hence $g|_{U \times A \times A} \in \mathcal{M}^k$. Furthermore $\sigma^k(g|_{U \times A \times A}) = (f_0, f_1, \ldots, f_k)$, so σ^k is surjective.

For all $k, l \in \mathbb{N}$, $k > l$, $\mathcal{G}_0^k$ has a natural 'action' [25] on $\mathcal{M}_0^l$ given by (use surjectivity of ρ^k and σ^l)

$$\begin{aligned}
&(D_0\phi, D_0^2\phi, \ldots, D_0^k\phi) \cdot (g_0, D_0 g, \ldots, D_0^l g) \\
&:= \Big((u_1, u_2) \mapsto g_0(D_0\phi^{-1}(u_1), D_0\phi^{-1}(u_2)), \\
&\quad (u_1, u_2, u_3) \mapsto D_0 g(D_0\phi^{-1}(u_3))(D_0\phi^{-1}(u_1), D_0\phi^{-1}(u_2)) \\
&\qquad + g_0(D_0^2\phi^{-1}(u_1, u_3), D_0\phi^{-1}(u_2)) + g_0(D_0\phi^{-1}(u_1), D_0^2\phi^{-1}(u_2, u_3)), \ldots\Big),
\end{aligned}$$

where all higher order terms can be obtained by writing out $D_0^m(\phi_* g)$ for $0 \leq m \leq l$ using Equation (5.4) and Equation (6.17). Note that with this definition, for $\phi \in \mathcal{G}^k$ and $g \in \mathcal{M}^l$ for which $\phi \cdot g$ is defined, we have

$$\sigma^l(\phi \cdot g) = \sigma^l(\phi_* g) = \rho^k(\phi) \cdot \sigma^l(g),$$

So in particular if $\chi\,\phi$ is defined in $\mathcal{G}^k$, $\rho^k(\chi) \cdot (\rho^k(\phi) \cdot \sigma^l(g)) = \rho^k(\chi) \cdot \sigma^l(\phi_* g) = \rho^k(\chi_*(\phi_* g)) = \rho^k((\chi \circ \phi)_* g) = \rho^k(\chi\,\phi) \cdot \sigma^l(g)$.

Now we will create a new container for sequences of the form $(g_0, R_0^g, \nabla R_0^g, \ldots)$. Let $k \in \mathbb{N}$.

Define $\mathcal{R}_0^{k,0}$ to be the collection of all $f : A^4 \to \mathbb{K}$ Ⓒ 4-① $/\mathbb{K}$ satisfying $f(a_1, a_2, a_3, a_4) = -f(a_2, a_1, a_3, a_4)$, $f(a_1, a_2, a_3, a_4) = -f(a_1, a_2, a_4, a_3)$, $f(a_1, a_2, a_3, a_4) = f(a_3, a_4, a_1, a_2)$, and $f(a_1, a_2, a_3, a_4) + f(a_1, a_4, a_2, a_3) + f(a_1, a_3, a_4, a_2) = 0$. For $l > 0$ define $\mathcal{R}_0^{k,l}$ to consist of all $f : A^{l+4} \to \mathbb{K}$ Ⓒ $(l+4)$-① $/\mathbb{K}$ that satisfy the same relations for their last four variables as the maps in $\mathcal{R}_0^{k,1}$. Then we define

$$\mathcal{R}_0^k := \mathcal{M}_0^{k,0} \cup \Big(\bigcup_{l=0}^{k-2} \mathcal{R}_0^{k,l}\Big).$$

[25] Not truly a group action, as $\mathcal{G}_0^k$ is not a group, but really a pairing $\mathcal{G}_0^k \times \mathcal{M}_0^l \to \mathcal{M}_0^l$: $(a, b) \mapsto a \cdot b$ which satisfies $a \cdot (b \cdot c) = (a\,b) \cdot c$.

Between $\mathcal{M}_0^k$ and $\mathcal{R}_0^k$ we can define the *Christoffel map* (which is well-defined by Equation (6.13), of which the relations are preserved by Equation (6.15))

$$\mathrm{Chri}^k : \mathcal{M}_0^k \to \mathcal{R}_0^k : (g_0, D_0 g, D_0^2 g, \dots, D_0^k g) \mapsto (g_0, R_0^g, \nabla R_0^g, \dots, \nabla^{k-2} R_0^g), \tag{6.37}$$

where we calculate R_0^g with Equation (6.11) and Γ_0^g using Equation (6.4) and Equation (6.7). Terms $\nabla^l R_0^g$ for $l \geq 1$ may inductively be calculated using Equation (6.15) and expressed entirely in terms of $D_0^m g$, where $D^l R_0^g$ is calculated with Theorem (5.1.8), Equation (5.4), and Equation (6.1) (which gives $D^l \Gamma_0^g$ in terms of derivatives $D_0^m g$ and Γ_0^g itself).

The group $\mathcal{G}_0^1$ of all c 1 bijective mappings $A \to A$ has a natural action on $\mathcal{R}_0^k$ given for all $\phi \in \mathcal{G}_0^1$ and $(f_2, f_4, \dots, f_{k+2}) \in \mathcal{R}_0^k$ by

$$\begin{aligned}
&\phi \cdot (f_2, f_4, f_5, \dots, f_{k+2}) \\
&:= \Big((u_1, u_2) \mapsto f_2(\phi^{-1}(u_1), \phi^{-1}(u_2)), \\
&\quad (u_1, u_2, u_3, u_4) \mapsto f_4(\phi^{-1}(u_1), \phi^{-1}(u_2), \phi^{-1}(u_3), \phi^{-1}(u_4)), \\
&\quad \dots, \\
&\quad (u_1, \dots, u_{k+2}) \mapsto f_{k+2}(\phi^{-1}(u_1), \dots, \phi^{-1}(u_{k+2})) \Big).
\end{aligned}$$

It can directly be verified that $\chi \cdot (\phi \cdot (f_2, f_4, \dots, f_{k+2})) = (\chi\, \phi) \cdot (f_2, f_4, \dots, f_{k+2})$ and $\mathrm{id}_A \cdot (f_2, f_4, \dots, f_{k+2}) = (f_2, f_4, \dots, f_{k+2})$, so this really is a group action.

Note that for any $\phi \in \mathcal{G}^k$ and $g \in \mathcal{M}^l$, $k > l$, $\phi \cdot g$ defined, we have with this definition that (using Lemma (5.1.15) and Theorem (6.2.3))

$$\begin{aligned}
&(\pi_1^k \circ \rho^k)(\phi) \cdot (\mathrm{Chri}^l \circ \sigma^l)(g) \\
&= D_0\phi \cdot (g_0, R_0^g, \nabla R_0^g, \dots) \\
&= \Big((u_1, u_2) \mapsto g_0([D_0\phi]^{-1}(u_1), [D_0\phi]^{-1}(u_2)), \\
&\quad (u_1, \dots, u_4) \mapsto R_0^g([D_0\phi]^{-1}(u_1), \dots, [D_0\phi]^{-1}(u_4)), \dots \Big) \\
&= \Big((u_1, u_2) \mapsto g_{\phi^{-1}(0)}(D_0\phi^{-1}(u_1), D_0\phi^{-1}(u_2)), \\
&\quad (u_1, \dots, u_4) \mapsto R^g_{\phi^{-1}(0)}(D_0\phi^{-1}(u_1), \dots, D_0\phi^{-1}(u_4)), \dots \Big) \\
&= \Big((\phi_* g)_0, R_0^{\phi_* g}, \dots \Big) \\
&= \mathrm{Chri}^l(\sigma^l(\phi_* g)) \\
&= (\mathrm{Chri}^l \circ \sigma^l)(\phi \cdot g).
\end{aligned}$$

This shows us that the action of $\mathcal{G}_0^1$ on $\mathcal{R}_0^l$ is compatible with the earlier actions we established.

To summarise all of the above, for $k, l \in \mathbb{N}$, $k > l$, we have

$$\phi \longmapsto (D_0\phi, D_0^2\phi, \ldots) \longmapsto D_0\phi \tag{6.38}$$

$$\begin{tikzcd}
\mathcal{G}^k \arrow[r, two heads, "\rho^k"] \arrow[d, dotted] & \mathcal{G}_0^k \arrow[r, two heads, "\pi_1^k"] \arrow[d, dotted] & \mathcal{G}_0^1 \arrow[d, dotted] \\
\mathcal{M}^l \arrow[r, two heads, "\sigma^l"'] & \mathcal{M}_0^l \arrow[r, "\mathrm{Chri}^l"'] & \mathcal{R}_0^l
\end{tikzcd}$$

$$g \longmapsto (g_0, D_0 g, \ldots) \longmapsto (g_0, R_0^g, \ldots).$$

Here the dotted arrows denote the actions induced by the pushforward $(\phi, g) \mapsto \phi_* g$. Furthermore, the actions and multiplication maps in the diagram are compatible whenever they are defined:

$$\begin{aligned}
\rho^k(\chi\,\phi) &= \rho^k(\chi)\,\rho^k(\phi) \\
\sigma^l(\phi \cdot g) &= \rho^k(\phi) \cdot \sigma^l(g) \\
(\pi_1^k \circ \rho^k)(\chi\,\phi) &= (\pi_1^k \circ \rho^k)(\chi)\,(\pi_1^k \circ \rho^k)(\phi) \\
(\mathrm{Chri}^l \circ \sigma^l)(\phi \cdot g) &= (\pi_1^k \circ \rho^k)(\phi) \cdot (\mathrm{Chri}^l \circ \sigma^l)(g).
\end{aligned} \tag{6.39}$$

However, we still have not used Theorem (6.6.1). We therefore define the collection of C^k metrics that are simple at 0 by

$$\mathcal{S}^k := \{g \in \mathcal{M}^k |\ g \text{ is a simple metric at } 0\ \} \subseteq \mathcal{M}^k.$$

Define for $k, l \in \mathbb{N}$ furthermore $\mathcal{S}_0^{k,0} := \mathcal{M}_0^{k,0}$, $\mathcal{S}_0^{k,1} := \{0\}$ as the set containing only the zero mapping $0 : A^3 \to A$ (inspired by Equation (6.26)), and for $l > 1$, $\mathcal{S}_0^{k,l}$ to consist of all $f \in \mathcal{M}_0^{k,l}$ which in addition satisfy (Equation (6.27))

$$\begin{aligned}
f(a_1, a_2, \ldots, a_l, a_{l+1}, a_{l+2}) + f(a_{l+1}, a_1, \ldots, a_{l-1}, a_l, a_{l+2}) + \ldots \\
+ f(a_2, a_3, \ldots, a_{l+1}, a_1, a_{l+2}) = 0
\end{aligned}$$

for all $a_1, \ldots, a_{l+2} \in A$. Then we define

$$\mathcal{S}_0^k := \bigcup_{l=0}^{k} \mathcal{S}_0^{k,l} \subseteq \mathcal{M}_0^k.$$

By Lemma (6.4.4) we have that for any $g \in \mathcal{S}^{k+1}$, $\sigma^k(g) \in \mathcal{S}_0^k$. [26] We see from the derivation of Equation (6.29) that Chri^k is injective when restricted to $\mathcal{S}_0^k$: all $(g_0, D_0 g, \ldots, D_0^k g) \in \mathcal{S}_0^k \subseteq \mathcal{M}_0^k$ satisfy Equation (6.26) and Equation (6.27) which were all that was necessary to derive Equation (6.29).

So Equation (6.29) and Theorem (6.4.6) can in a sense be seen as expressesions of the fact that the map Chri^k is injective when restricted to $\mathcal{S}_0^k$.

[26] It is unclear whether or not $\sigma^k(\mathcal{S}^{k+1}) = \mathcal{S}_0^k$ however.

Whether or not Chri^k is surjective is a harder problem: Equation (6.29) gives us a way to construct a Taylor sequence of a metric up to order k, given a sequence $(f_2, f_4, f_5, \ldots, f_{k+2}) \in \mathcal{R}_0^k$, but whether or not this Taylor sequence again yields $(f_2, f_4, f_5, \ldots, f_{k+2})$ under Chri^k is unclear, because Equation (6.29) and Chri^k both become very complicated at higher orders.

By Lemma (6.4.4) we furthermore know that for any $\phi \in \mathcal{G}_0^1$ and $g \in \mathcal{S}^k$, we have that $\phi \cdot g \in \mathcal{S}^k$ if $k \geq 2$, after suitably restricting ϕ. Likewise, all $\phi \in \mathcal{G}_0^1$ preserve the relations of Equation (6.26) and Equation (6.27), so $\phi \cdot \mathcal{S}_0^k \subseteq \mathcal{S}_0^k$. Hence for all $k \in \mathbb{N}$

$$\begin{tikzcd} \mathcal{G}_0^1 \arrow[r, tail, two heads] \arrow[d, dotted] & \mathcal{G}_0^1 \arrow[r, tail, two heads] \arrow[d, dotted] & \mathcal{G}_0^1 \arrow[d, dotted] \\ \mathcal{S}^{k+1} \arrow[r, "\sigma^k"'] & \mathcal{S}_0^k \arrow[r, tail, "\mathrm{Chri}^k"'] & \mathcal{R}_0^k. \end{tikzcd} \tag{6.40}$$

Theorem (6.5.1) states that for $k \geq 3$, and any $g \in \mathcal{M}^k$ there exists a $\phi \in \mathcal{G}^{k-1}$ such that $\phi{\cdot}g \in \mathcal{S}^{k-2}$ if g is restricted to a suitable open neighbourhood of 0.

Let $k \in \mathbb{N}$, $(g_0, D_0 g, \ldots, D_0^k g) \in \mathcal{M}_0^k$ be arbitrary, then we may assume that $g \in \mathcal{M}^\infty$ by the proof of surjectivity of σ^k. Hence by Theorem (6.5.1) there exists a $\phi \in \mathcal{G}^\infty$ such that $(D_0\phi, D_0^2\phi, \ldots, D_0^{k+1}\phi) \cdot (g_0, D_0 g, \ldots, D_0^k g) = \sigma^k(\phi \cdot g) \in \sigma^k(\mathcal{S}^{k+1}) \subseteq \mathcal{S}_0^k$. So for any $(g_0, D_0 g, \ldots) \in \mathcal{M}_0^k$ there exists a $(D_0\phi, D_0^2\phi, \ldots, D_0^{k+1}\phi) \in \mathcal{G}_0^{k+1}$ such that $(D_0\phi, D_0^2\phi, \ldots) \cdot (g_0, D_0 g, \ldots) \in \mathcal{S}_0^k$.

Therefore Theorem (6.5.1) expresses in this setting the fact that for any $k \in \mathbb{N}$, the 'orbit' (as it is not really a group action) of any Taylor sequence in $\mathcal{M}_0^k$ under $\mathcal{G}_0^{k+1}$ necessarily intersects $\mathcal{S}_0^k$. So we can go from Equation (6.38) to Equation (6.40), while staying in the same orbit in $\mathcal{M}_0^k$.

The mayor advantage of this, is that the action of $\mathcal{G}_0^1$ on $\mathcal{S}_0^k$ and $\mathcal{R}_0^k$ is much less complicated than the action of $\mathcal{G}_0^{k+1}$ on $\mathcal{M}_0^k$ and therefore a much more convenient setting in which to investigate whether or not two elements lie in the same orbit.

This is another way, compared to Theorem (6.6.1), of looking at the original problem in the context of orbits and expressing Theorem (6.4.6) and Theorem (6.5.1) in terms of this new formulation.

Chapter 7

Conclusion

In the first chapters of this thesis we discussed topology (Chapter 2), algebra and group theory (Chapter 3), algebra and topology combined (Chapter 4), and analysis (Chapter 5) to develop the theory necessary to generalise Christoffel's article, [Chr1869], as much as possible. The many interesting results found in these chapters are listed in Section 1.2.

Christoffel's article deals with the question of, given two metrics, whether or not we can find a coordinate transformation that transforms these metrics into each other, as per Equation (6.2). In [Chr1869] the consequences of Equation (6.2) are investigated to find necessary conditions, for such a coordinate transformation to exist. These conditions have all been generalised from the finite dimensional space $\mathbb{R}^k$ to TVS /$\mathbb{K}$ T2 LC (that is, locally convex Hausdorff topological vector spaces over either $\mathbb{R}$ or $\mathbb{C}$) in Theorem (6.2.1), Theorem (6.2.2), and Theorem (6.2.3). In these theorems we find that if such a coordinate transformation exists, we obtain a chain of transformation equations between tensors which are all necessarily satisfied by the transformation.[1] Furthermore, these covariant tensors can directly be expressed in terms of the derivatives of the metrics with which they are associated.

Then we digressed from [Chr1869] in Section 6.3 at the point where we try to find a way back from this chain of equations to a coordinate transformation relating the metrics. Because considering all coordinate transformations in their entirety would make this extremely complicated, we fixed two points and approximated both metrics by their Taylor approximations near these points. In Theorem (6.4.6) we then found that the Taylor approximations to two given metrics agree at a certain point, precisely when the sequence of tensors from [Chr1869] agrees at this point, provided that we are working in a UC (uniformly complete) space, and both the metrics are simple[2] at this point. We continued with Theorem (6.5.1) which shows that if we are working in a Ba (Banach) space, then for any metric and any fixed point at which the metric is defined, there is a coordinate transformation such that the transformed metric is simple at that point. This in turn lead to Theorem (6.6.1) where we show for two fixed

[1] More precisely, we find equations between the curvature tensors of both metrics and their covariant derivatives.

[2] Expressed in Riemann normal coordinates.

points, that we can find coordinate transformations for both metrics, such that the transformed metrics at these points have an equal Taylor expansion, if and only if there is a continuous linear bijective map which relates the tensors from [Chr1869] at these points.

While this does not provide the coordinate transformation itself, it does simplify the question of whether or not there can be a coordinate transformation relating the two fixed points in question, which preserves the Taylor sequences of both metrics up to a certain order, by rephrasing this question entirely in terms of the tensors from [Chr1869]. It furthermore ensures that at these points, we need not consider the entire Taylor expansion of the coordinate transformation, but only its linear part.

At the end of Section 6.3 and continued in Section 6.6 we rephrase the problem in terms of an action of the collection of diffeomorphisms on the collection of metrics (such that two metrics can be transformed into each other via a coordinate transformation if and only if they lie in the same orbit under this action). Here we see that the tensors from [Chr1869] give us a map, given by Equation (6.37), called the *Christoffel map* which factors nicely through the action of the diffeomorphisms (Equation (6.39)), and is even injective on a certain subset of Taylor sequences of simple metrics by Equation (6.29).

This raises some interesting questions.

- At the end of Section 6.6 we found that the Christoffel map Chri^k is injective when restricted to $\mathcal{S}_0^k$. This set $\mathcal{S}_0^k$ contains $\sigma^k(\mathcal{S}^{k+1})$, the Taylor sequences of all C^{k+1} metrics that are simple at 0. We can therefore ask ourselves whether $\sigma^k(\mathcal{S}^k)$ or $\sigma^k(\mathcal{S}^{k+1})$ equals the entire $\mathcal{S}_0^k$.

- Another immediate question is what the set $\mathrm{Chri}^k(\mathcal{S}_0^k) \subseteq \mathcal{R}_0^k$ looks like. Is Chri^k defined in Equation (6.37) surjective? Or in other words, what conditions must the multilinear maps from a given sequence $(g_0, R_0^g, \nabla R_0^g, \ldots)$ satisfy to actually be of this form for a certain metric g?

- And finally, would there be a practical way to employ Theorem (6.6.1) in finding the sought coordinate transformation, or is it only useful in eliminating points that may not be related by any coordinate transformation?

These questions mark the end of this thesis. I would like to thank the reader for taking the time to read it, and I hope to have given an interesting new perspective on Christoffel's article,

Bas Fagginger Auer.

Chapter 8

Translation

This is a annotated, direct translation of Elwin Bruno Christoffel's article, [Chr1869]. All footnotes have been added by the author of this thesis, except for the references to the articles of Lamé and Riemann.

About the transformation of homogeneous differential expressions of second degree

(Mr. *E. B. Christoffel* from Zürich, translated by *B. O. Fagginger Auer*)

1.

Consider the differential expression

$$F = \sum \omega_{ik}\, \partial x_i\, \partial x_k; \qquad i, k = 1, 2, \ldots, n,$$

where the coefficients ω are arbitrary functions of the independent variables x_1, x_2, $\ldots$, x_n. If these are independent functions of new variables x'_1, x'_2, $\ldots$, x'_n then F will transform into a new differential expression

$$F' = \sum \omega'_{ik}\, \partial x'_i\, \partial x'_k,$$

which is equal to the original expression.

If on the other hand two differential expressions F and F' are given, one can ask under which conditions these may be transformed into each other and which substitution of variables will yield this transformation.

For the investigation of this question we define the determinants of the coefficients of F, F', and the substitution itself as

$$\sum \pm \frac{\partial x_1}{\partial x'_1} \frac{\partial x_2}{\partial x'_2} \cdots \frac{\partial x_n}{\partial x'_n} = r,$$
$$\sum \pm \omega_{11}\, \omega_{22} \ldots \omega_{nn} = E,$$
$$\sum \pm \omega'_{11}\, \omega'_{22} \ldots \omega'_{nn} = E'.$$

From the theory of algebraic invariants it is known that there is precisely one condition

$$E' = r^2\, E$$

for our sought after transformation, which would be sufficient for the considered case, provided the coefficients ω of F and elements of r are constant. In the actual, more general, case we are considering, additional conditions need to be met, as not all systems of linear and homogeneous functions of $\partial x'_1$, $\partial x'_2$, $\ldots$, $\partial x'_n$ inserted in place of ∂x_1, ∂x_2, $\ldots$, ∂x_n to transform F to F' will solve the posed problem, without satisfying integrability conditions [1] that make the expressions for ∂x_1, ∂x_2, $\ldots$, ∂x_n full-fledged differentials.

If F and F' are homogeneous differential expressions of general, but equal degree, the theory of algebraic invariants yields equations which, in the case that this degree is strictly greater than 2, completely determine the values of the coefficients necessary to transform F into F'. However, these coefficients still need to be subjected to the necessary integrability conditions, which involves both the direct and inverse substitution, see section 10..[2]

[1] See the integrability conditions (B.) in section 8..

[2] *"[...], wobei die Eigenschaft der zugehörigen Formen, unmittelbar nicht die directe, sondern die transponirte Substitution zu liefern, wesentlich in Betracht kommt (vergl. art. 10)."*

These kind of simplifications do not occur for homogeneous differential expressions of second degree, as the algebraic conditions yield only one invariant and one corresponding form. Because of this, we will introduce aids in the following sections, with which this important case can be treated. Finally, it should be remarked that for the case where $F = \partial x_1^2 + \partial x_2^2 + \partial x_3^2$, there exists an extensive work of Mr. *Lamé* (Théorie des coordonnées curvilignes), treating our present problem in the context of the unwrapping of curved planes.

2.

During the investigation of the conditions that are both necessary and sufficient for the equation

$$(1.) \qquad \sum \omega_{ik}\,\partial x_i\,\partial x_k = \sum \omega'_{ik}\,\partial x'_i\,\partial x'_k$$

to hold, we restrict ourselves to the case where both the determinants E and E' of these differential expressions are not identically equal to zero.[3] The new variables x' will be assumed to be independent and their differentials constant.

If one would replace each $\partial x'_i$ by $\partial x'_i + \delta x'_i$, where the differentials $\delta x'_i$ correspond to increases of the original variables x by δx, then in (1.), ∂x would go to $\partial x + \delta x$. Expanding the products on both sides of (1.) we obtain[4]

$$(2.) \qquad \sum \omega_{ik}\,\partial x_i\,\delta x_k = \sum \omega'_{ik}\,\partial x'_i\,\delta x'_k.$$

Comparing the coefficients for $\partial x'_g$, we find

$$\sum_{ik} \omega_{ik}\,\frac{\partial x_i}{\partial x'_g}\,\delta x_k = \sum_k \omega'_{gk}\,\delta x'_k,$$

and from this

$$(3.) \qquad \delta x'_h = \sum_{gik} \omega_{ik}\,\frac{E'_{gh}}{E'}\,\frac{\partial x_i}{\partial x'_g}\,\delta x_k,$$

where E'_{gh} is the minor of ω'_{gh}.[5]

In equation (2.) we now increase each x' by its differential $\partial x'$ and in (1.) each x' by $\delta x'$, which does not change the differentials on the right hand side;

[3] It is also implicitly assumed that $\omega_{ik} = \omega_{ki}$ and $\omega'_{ik} = \omega'_{ki}$.

[4] For F we have $\sum \omega_{ik}\,\partial x_i\,\partial x_k \to \sum \omega_{ik}(\partial x_i + \delta x_i)(\partial x_k + \delta x_k) = \sum \omega_{ik}\,\partial x_i\,\partial x_k + \sum \omega_{ik}\,\delta x_i\,\delta x_k + \sum \omega_{ik}\,\partial x_i\,\delta x_k + \sum \omega_{ik}\,\delta x_i\,\partial x_k$, do the same for F' and use equation (1.) and symmetry of ω and ω' to obtain (2.).

[5] So E'_{gh} is the determinant of the matrix obtained from deleting the g-th row and h-th column of the matrix with coefficients $\left(\omega'_{ik}\right)_{i,k=1,\ldots,n}$. By Cramer's rule and assumed invertibility of ω', we have therefore $\sum_h (E'_{gh}/E')\omega'_{hk} = \delta_{gk}$, which immediately yields equation (3.) by multiplying the right hand side of the equation above (3.) by E'_{gh}/E' and summing over g.

it then follows that:[6]

$$\sum \omega_{ik}\, \partial^2 x_i\, \delta x_k + \sum \partial\omega_{ik}\, \partial x_i\, \delta x_k + \sum \omega_{ik}\, \partial x_i\, \partial\delta x_k = \sum \partial\omega'_{ik}\, \partial x'_i\, \delta x'_k,$$
$$\sum \delta\omega_{ik}\, \partial x_i\, \partial x_k + 2\sum \omega_{ik}\, \partial x_i\, \delta\partial x_k = \sum \delta\omega'_{ik}\, \partial x'_i\, \partial x'_k.$$

Dividing the second equation by 2 and subtracting it from the first we find

$$\sum \omega_{ik}\, \partial^2 x_i\, \delta x_k + \sum \partial\omega_{ik}\, \partial x_i\, \delta x_k - \frac{1}{2}\sum \delta\omega_{ik}\, \partial x_i\, \partial x_k$$
$$= \sum \partial\omega'_{ik}\, \partial x'_i\, \delta x'_k - \frac{1}{2}\sum \partial\omega'_{ik}\, \partial x'_i\, \partial x'_k.$$

To avoid cluttering our formulae we define

$$(4.) \qquad \frac{1}{2}\left[\frac{\partial\omega_{gk}}{\partial x_h} + \frac{\partial\omega_{hk}}{\partial x_g} - \frac{\partial\omega_{gh}}{\partial x_k}\right] = \begin{bmatrix} gh \\ k \end{bmatrix}$$

from which

$$(5.) \qquad \begin{bmatrix} hg \\ k \end{bmatrix} = \begin{bmatrix} gh \\ k \end{bmatrix} \quad \text{and} \quad \frac{\partial\omega_{hk}}{\partial x_g} = \begin{bmatrix} gh \\ k \end{bmatrix} + \begin{bmatrix} gk \\ h \end{bmatrix}$$

follow, and if we use the same notation for the transformed differential expression, we find[7]

$$(6.) \qquad \sum_{ik} \omega_{ik}\, \partial^2 x_i\, \delta x_k + \sum_{ikl} \begin{bmatrix} il \\ k \end{bmatrix} \partial x_i\, \partial x_l\, \delta x_k = \sum_{\alpha\beta h} \begin{bmatrix} \alpha\beta \\ h \end{bmatrix}' \partial x'_a\, \partial x'_\beta\, \delta x'_h.$$

Substituting $\delta x'_h$ by its value from (3.) causes the right hand side to change to

$$\sum_{\alpha\beta h} \begin{bmatrix} \alpha\beta \\ h \end{bmatrix}' \partial x'_\alpha\, \partial x'_\beta \sum_{gik} \omega_{ik}\, \frac{E'_{gh}}{E'}\, \frac{\partial x_i}{\partial x'_g}\, \delta x_k.$$

By equating coefficients of δx_k we find

$$\sum_i \omega_{ik}\, \partial^2 x_i + \sum_{il} \begin{bmatrix} il \\ k \end{bmatrix} \partial x_i\, \partial x_l = \sum_{gi\alpha\beta} \omega_{ik}\, \frac{\partial x_i}{\partial x'_g}\, \partial x'_\alpha\, \partial x'_\beta \sum_h \begin{bmatrix} \alpha\beta \\ h \end{bmatrix}' \frac{E'_{gh}}{E'}.$$

[6] We are making a first order Taylor approximation in both (1.) and (2.) by letting $x'_i \to x'_i + \delta x'_i$ and $x'_i \to x'_i + \partial x'_i$ respectively (so $\omega'(x'+\delta x')_{ik} = \omega'(x')_{ik} + \delta\omega'_{ik}$). Note that all ω_{ik}, ∂x_i, δx_i are functions of x' via the dependence $x = x(x')$. From equation (2.) we obtain under $x'_i \to x'_i + \partial x_i$ that for the right hand side $\sum \omega'_{ik}\, \partial x'_i\, \delta x'_k \to \sum \omega'_{ik}\, \partial x'_i\, \delta x'_k + \sum \partial\omega'_{ik}\, \partial x'_i\, \delta x'_k$, while for the left hand side (as the x_i depend on the x'_i) $\sum \omega_{ik}\, \partial x_i\, \delta x_k \to \sum \omega_{ik}\, \partial x_i\, \delta x_k + \sum \partial\omega_{ik}\, \partial x_i\, \delta x_k + \sum \omega_{ik}\, \partial^2 x_i\, \delta x_k + \sum \omega_{ik}\, \partial x_i\, \partial\delta x_k$. Now equating both sides and using (2.) we obtain the first equation. The second equation is found by doing the same for equation (1.) and using the symmetry of ω.

[7] Using equation (5.) we find

$$\sum \partial\omega_{ik}\, \partial x_i\, \delta x_k - \frac{1}{2}\sum \delta\omega_{ik}\, \partial x_i\, \delta x_k$$
$$= \sum \left(\begin{bmatrix} ji \\ k \end{bmatrix} + \begin{bmatrix} jk \\ i \end{bmatrix} \right) \partial x_j\, \partial x_i\, \delta x_k - \frac{1}{2}\sum \left(\begin{bmatrix} ji \\ k \end{bmatrix} + \begin{bmatrix} jk \\ i \end{bmatrix} \right) \delta x_j\, \partial x_i\, \partial x_k$$
$$= \sum \left(\begin{bmatrix} ji \\ k \end{bmatrix} + \begin{bmatrix} jk \\ i \end{bmatrix} - \frac{1}{2}\begin{bmatrix} ki \\ j \end{bmatrix} - \frac{1}{2}\begin{bmatrix} jk \\ i \end{bmatrix} \right) \partial x_j\, \partial x_i\, \delta x_k.$$

Now use the fact that this expression is symmetrical in j and i, together with equation (5.) to obtain (6.).

We now multiply this equation by E_{rk}/E and sum over k. If we define[8]

$$(7.)\qquad \sum_k \begin{bmatrix} il \\ k \end{bmatrix} \frac{E_{rk}}{E} = \begin{Bmatrix} il \\ r \end{Bmatrix}$$

from which

$$(8.)\qquad \begin{Bmatrix} il \\ r \end{Bmatrix} = \begin{Bmatrix} li \\ r \end{Bmatrix}$$

follows, it is found that

$$\partial^2 x_r + \sum_{il} \begin{Bmatrix} il \\ r \end{Bmatrix} \partial x_i \, \partial x_l = \sum_{\alpha\beta g} \begin{Bmatrix} \alpha\beta \\ g \end{Bmatrix}' \frac{\partial x_r}{\partial x'_g} \partial x'_\alpha \, \partial x'_\beta,$$

which on the other hand implies equation (6.)[9], therefore, for all α, β and r:

$$(9.)\qquad \frac{\partial^2 x_r}{\partial x'_\alpha \, \partial x'_\beta} + \sum_{ik} \begin{Bmatrix} ik \\ r \end{Bmatrix} \frac{\partial x_i}{\partial x'_\alpha} \frac{\partial x_k}{\partial x'_\beta} = \sum_\lambda \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}' \frac{\partial x_r}{\partial x'_\lambda}.$$

This equation yields $n(n+1)/2$ equations for each r, so in total a system of $n^2(n+1)/2$ partial differential equations for the sought after substitution of variables. If these are satisfied, then so are the integrability conditions for the linear expression in the new differentials. [10]

These equations simplify considerably if all coefficients ω of the original form are constant. This, by (7.) and (5.), is true if and only if all expressions $\begin{Bmatrix} ik \\ r \end{Bmatrix}$ disappear. In this case we retain a linear system of partial differential equations for all variables. This result was also obtained by Mr. *Lamé* in the aforementioned case he was considering.

3.

If it is impossible to express the original variables x as functions of x' such that (9.) is satisfied, then the transformation from F to F' is not possible, as (9.) is a direct consequence of the existence of such a change of variables.

If on the other hand the equations (9.) satisfied, then one wonders in how far this leads back to equation (1.), that is $F = F'$. To investigate this, we insert the found solution to (9.) into F, from which follows that

$$F = \sum \omega''_{ik} \, \partial x'_i \, \partial x'_k = F''.$$

[8] These are the actual *Christoffel symbols*: for a pseudo-Riemannian metric g on a smooth manifold M we can write g in local coordinate patches as $g(x) = \sum_{a,b=1}^m g_{ab}(x)\, \mathrm{d}x^a \otimes \mathrm{d}x^b$, where $g_{ab}(x) = g_{ba}(x)$ is symmetric and the matrix $\big(g_{ab}(x)\big)_{a,b=1,\dots,m}$ is invertible. Denoting the inverse matrix by $\big(g^{ab}\big)_{a,b=1,\dots,m}$ we find that if we put $g_{ab} = \omega_{ab}$, such that $g = F$, by Cramer's rule $g^{ab} = E_{ab}/E$, so by equation (4.) and (7.) $\begin{Bmatrix} ab \\ c \end{Bmatrix} = \sum_d \begin{bmatrix} ab \\ d \end{bmatrix} \frac{E_{d,c}}{E} = \sum_d \frac{1}{2} \left[\frac{\partial \omega_{ad}}{\partial x_b} + \frac{\partial \omega_{bd}}{\partial x_a} - \frac{\partial \omega_{ab}}{\partial x_d} \right] \frac{E_{d,c}}{E} = \sum_d \frac{1}{2} g^{cd} \left(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab} \right) = \Gamma^c_{ab}$.

[9] The equations are equivalent as ω is invertible.

[10] By equation (8.) we have symmetry in α and β of (9.).

The important thing now is what connections there exist between the respective coefficients of F'' and F'.

As (9.) is nothing but a rewriting of equation (6.), both equations are equivalent. If we now, instead of starting with $F = F'$, begin by setting $F = F''$ and follow the same reasoning as in section 2, then instead of the current right hand side of (6.), we would obtain the new expression

$$\sum_{\alpha\beta h} \begin{bmatrix} \alpha\beta \\ h \end{bmatrix}'' \partial x'_\alpha \, \partial x'_\beta \, \delta x'_h,$$

and this must equal the original left hand side of (6.). Therefore we find

$$\begin{bmatrix} \alpha\beta \\ h \end{bmatrix}'' = \begin{bmatrix} \alpha\beta \\ h \end{bmatrix}',$$

which with (5.) gives

$$\frac{\partial \omega''_{hk}}{\partial x'_g} = \frac{\partial \omega'_{hk}}{\partial x'_g}$$

for all values of g, h, and k. Therefore the coefficients of F'' and F' can only differ by additive constants. To make these constants disappear, it is sufficient that $F'' = F'$ at a certain point, that is, that the solution to (9.) satisfies the transformation relations following from equation (1.):

$$(10.) \qquad \sum_{ik} \omega_{ik} \frac{\partial x_i}{\partial x'_\alpha} \frac{\partial x_k}{\partial x'_\beta} = \omega'_{\alpha\beta}$$

for a certain collection of values for the new variables. We therefore arrive at the statement:

Whenever it is possible to express the original variables $x_1, x_2, \ldots, x_n$ *in terms of the new variables* $x'_1, x'_2, \ldots, x'_n$*, such that the system of equations* (9.) *is satisfied, together with the initial condition that their first derivatives for a certain value of the new variables satisfy the transformation relations following from* $F = F'$*, then* $F = F'$ *for all possible values of the new variables.*

From this it follows that the transformation relations contained in (1.), except from the required initial conditions, are made redundant by (9.).

The equations contained in (9.) *are, provided that the initial conditions are satisfied at a certain point, both necessary and sufficient for existence of a transformation of* F *into* F' *and completely replaces the algebraic and integrability conditions stemming from equation* (1.).

4.

For the possibility of the equations (9.) being satisfied, new integrability conditions are necessary which we will now derive. To also show that these conditions do not again imply (9.), we will consider the following, more general equations instead of (9.)

$$(9'.) \qquad \begin{cases} \dfrac{\partial^2 x_r}{\partial x'_\alpha \partial x'_\beta} + \sum\limits_{gh} \begin{Bmatrix} gh \\ r \end{Bmatrix} u^g_\alpha u^h_\beta = \sum\limits_{\lambda} \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}' u^r_\lambda + \begin{pmatrix} r \\ \alpha\beta \end{pmatrix} \\ \dfrac{\partial^2 x_r}{\partial x'_\alpha \partial x'_\gamma} + \sum\limits_{gh} \begin{Bmatrix} gh \\ r \end{Bmatrix} u^g_\alpha u^h_\gamma = \sum\limits_{\lambda} \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}' u^r_\lambda + \begin{pmatrix} r \\ \alpha\gamma \end{pmatrix} \end{cases}$$

where, as we will do in the following, denote the first derivatives

$$\frac{\partial x_i}{\partial x'_\alpha} = u^i_\alpha$$

while we use the usual notation for higher order derivatives. Note that we recover equation (9.) when we put $\begin{pmatrix} r \\ \alpha\beta \end{pmatrix}$ equal to 0.

We obtain the discussed integrability conditions when we differentiate the first equation of (9′.) to x'_γ, the second to x'_β and consider their difference, where the third derivative of x_r disappears. Also noting that the terms which contain a derivative with respect to both x'_β and x'_γ disappear[11], it follows:

$$\sum_{ghi} \left[\frac{\partial \begin{Bmatrix} gh \\ r \end{Bmatrix}}{\partial x_i} - \frac{\partial \begin{Bmatrix} gi \\ r \end{Bmatrix}}{\partial x_h} \right] u^g_\alpha u^h_\beta u^i_\gamma + \sum_{ph} \begin{Bmatrix} ph \\ r \end{Bmatrix} \frac{\partial^2 x_p}{\partial x'_\alpha \partial x'_\gamma} u^h_\beta - \sum_{pi} \begin{Bmatrix} pi \\ r \end{Bmatrix} \frac{\partial^2 x_p}{\partial x'_\alpha \partial x'_\beta} u^i_\gamma$$
$$= \sum_\lambda \left[\frac{\partial \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}'}{\partial x'_\gamma} - \frac{\partial \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}'}{\partial x'_\beta} \right] u^r_\lambda + \sum_\lambda \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}' \frac{\partial^2 x_r}{\partial x'_\lambda \partial x'_\gamma} - \sum_\lambda \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}' \frac{\partial^2 x_r}{\partial x'_\lambda \partial x'_\beta}$$
$$+ \frac{\partial \begin{pmatrix} r \\ \alpha\beta \end{pmatrix}}{\partial x'_\gamma} - \frac{\partial \begin{pmatrix} r \\ \alpha\gamma \end{pmatrix}}{\partial x'_\beta}.$$

From this it follows that the integrability conditions for (9′.) are precisely those for our original equations (9.), whenever the $\begin{pmatrix} r \\ \alpha\beta \end{pmatrix}$ satisfy the following equation

$$(11.) \quad \begin{cases} \displaystyle\sum_{ph} \begin{Bmatrix} ph \\ r \end{Bmatrix} \begin{pmatrix} p \\ \alpha\gamma \end{pmatrix} u^b_\beta - \sum_{pi} \begin{Bmatrix} pi \\ r \end{Bmatrix} \begin{pmatrix} p \\ \alpha\beta \end{pmatrix} u^i_\gamma \\ \displaystyle = \sum_\lambda \begin{Bmatrix} \alpha\beta \\ \gamma \end{Bmatrix}' \begin{pmatrix} r \\ \lambda\gamma \end{pmatrix} - \sum_\lambda \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}' \begin{pmatrix} r \\ \lambda\beta \end{pmatrix} + \frac{\partial \begin{pmatrix} r \\ \alpha\beta \end{pmatrix}}{\partial x'_\gamma} - \frac{\partial \begin{pmatrix} r \\ \alpha\gamma \end{pmatrix}}{\partial x'_\beta}, \end{cases}$$

which can easily be put in a more symmetric form.

From now on we will suppose that this condition is satisfied for all values of r, α, β, and γ. If we then substitute all second derivatives by their corresponding

[11] Symmetry of higher order derivatives.

expressions in (9′.), we obtain:

$$
\begin{aligned}
&\sum_{ghi}\left[\frac{\partial \begin{Bmatrix} gh \\ r \end{Bmatrix}}{\partial x_i} - \frac{\partial \begin{Bmatrix} gi \\ r \end{Bmatrix}}{\partial x_h}\right] u_\alpha^g u_\beta^h u_\gamma^i \\
&\qquad + \sum_{ph} \begin{Bmatrix} ph \\ r \end{Bmatrix} u_\beta^h \left(\sum_\lambda \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}' u_\lambda^p - \sum_{gi} \begin{Bmatrix} gi \\ p \end{Bmatrix} u_\alpha^g u_\gamma^i \right) \\
&\qquad - \sum_{ph} \begin{Bmatrix} ph \\ r \end{Bmatrix} u_\gamma^h \left(\sum_\lambda \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}' u_\lambda^p - \sum_{gi} \begin{Bmatrix} gi \\ p \end{Bmatrix} u_\alpha^g u_\beta^i \right) \\
&\qquad\qquad = \sum_\lambda \left[\frac{\partial \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}'}{\partial x'_\gamma} - \frac{\partial \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}'}{\partial x'_\beta}\right] u_\lambda^r \\
&\qquad + \sum_\lambda \begin{Bmatrix} \alpha\beta \\ \lambda \end{Bmatrix}' \left(\sum_\mu \begin{Bmatrix} \gamma\lambda \\ \mu \end{Bmatrix}' u_\mu^r - \sum_{pi} \begin{Bmatrix} pi \\ r \end{Bmatrix} u_\lambda^p u_\gamma^i \right) \\
&\qquad\qquad - \sum_\lambda \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}' \left(\sum_\mu \begin{Bmatrix} \beta\lambda \\ \mu \end{Bmatrix}' u_\mu^r - \sum_{pi} \begin{Bmatrix} pi \\ r \end{Bmatrix} u_\lambda^p u_\beta^i \right).
\end{aligned}
$$

Here all terms containing the product of two u's cancel; if we collect the remaining terms and swap λ and μ, it follows that:

$$
(12.)\quad \left\{
\begin{aligned}
&\sum_{ghi}\left(\frac{\partial \begin{Bmatrix} gh \\ r \end{Bmatrix}}{\partial x_i} - \frac{\partial \begin{Bmatrix} gi \\ r \end{Bmatrix}}{\partial x_h} + \sum_p \left[\begin{Bmatrix} gh \\ p \end{Bmatrix} \begin{Bmatrix} pi \\ r \end{Bmatrix} - \begin{Bmatrix} gi \\ p \end{Bmatrix} \begin{Bmatrix} ph \\ r \end{Bmatrix} \right] \right) u_\alpha^g u_\beta^h u_\gamma^i \\
&= \sum_\lambda \left(\frac{\partial \begin{Bmatrix} \alpha\beta \\ \gamma \end{Bmatrix}'}{\partial x'_\gamma} - \frac{\partial \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}'}{\partial x'_\beta} + \sum_\mu \left[\begin{Bmatrix} \alpha\beta \\ \mu \end{Bmatrix}' \begin{Bmatrix} \mu\gamma \\ \lambda \end{Bmatrix}' - \begin{Bmatrix} \alpha\gamma \\ \mu \end{Bmatrix}' \begin{Bmatrix} \mu\beta \\ \lambda \end{Bmatrix}' \right] \right) u_\lambda^r.
\end{aligned}
\right.
$$

We therefore obtain the following result.

The system of equations (12.) *where* r, α, β, *and* γ *range from* 1 *to* n, *contains the necessary integrability conditions of equation* (9.). *However, it cannot replace* (9.), *as it implies the more general system of equations* (9′.) *where the independent terms* $\begin{pmatrix} r \\ \alpha\beta \end{pmatrix}$ *satisfy* (11.).

5.

To continue our investigation, the equations contained in (12.) have to be replaced by a different, but equivalent system of equations. To this end we multiply (12.) by $\omega_{rk} u^k_\delta$ and sum over k and r. It is clear that the equations this yields are equivalent to (12.): we can multiply the equations we obtain by the above operations by $\frac{\partial x'_\delta}{\partial x_l} \frac{E_{\rho l}}{E}$, sum over δ and l, and finally set $\rho = r$ to again obtain (12.).

Performing these operations we obtain from (12.), using (10.)

$$\sum_{rk} \omega_{rk} u^r_\lambda u^k_\delta = \omega'_{\lambda\delta}$$

the following equation

$$\sum_{ghik} u^g_\alpha u^h_\beta u^i_\gamma u^k_\delta \sum_r \omega_{rk} \left(\frac{\partial \begin{Bmatrix} gh \\ r \end{Bmatrix}}{\partial x_i} - \frac{\partial \begin{Bmatrix} gi \\ r \end{Bmatrix}}{\partial x_h} + \sum_p \left[\begin{Bmatrix} gh \\ p \end{Bmatrix} \begin{Bmatrix} pi \\ r \end{Bmatrix} - \begin{Bmatrix} gi \\ p \end{Bmatrix} \begin{Bmatrix} ph \\ r \end{Bmatrix} \right] \right) u^g_\alpha u^h_\beta u^i_\gamma$$

$$= \sum_\lambda \omega'_{\lambda\delta} \left(\frac{\partial \begin{Bmatrix} \alpha\beta \\ \gamma \end{Bmatrix}'}{\partial x'_\gamma} - \frac{\partial \begin{Bmatrix} \alpha\gamma \\ \lambda \end{Bmatrix}'}{\partial x'_\beta} + \sum_\mu \left[\begin{Bmatrix} \alpha\beta \\ \mu \end{Bmatrix}' \begin{Bmatrix} \mu\gamma \\ \lambda \end{Bmatrix}' - \begin{Bmatrix} \alpha\gamma \\ \mu \end{Bmatrix}' \begin{Bmatrix} \mu\beta \\ \lambda \end{Bmatrix}' \right] \right)$$

such that the following manipulations will be the same for both sides of the equation. Now by equation (7.)

$$\sum_r \omega_{rk} \begin{Bmatrix} gh \\ r \end{Bmatrix} = \begin{bmatrix} gh \\ k \end{bmatrix},$$

hence carrying out the sum over r on the left hand side

$$= \frac{\partial \begin{bmatrix} gh \\ k \end{bmatrix}}{\partial x_i} - \sum_p \begin{Bmatrix} gh \\ p \end{Bmatrix} \frac{\partial \omega_{pk}}{\partial x_i} - \frac{\partial \begin{bmatrix} gi \\ k \end{bmatrix}}{\partial x_h} + \sum_p \begin{Bmatrix} gi \\ p \end{Bmatrix} \frac{\partial \omega_{pk}}{\partial x_h} + \sum_p \left(\begin{Bmatrix} gh \\ p \end{Bmatrix} \begin{bmatrix} pi \\ k \end{bmatrix} - \begin{Bmatrix} gi \\ p \end{Bmatrix} \begin{bmatrix} ph \\ k \end{bmatrix} \right),$$

where in both sums containing derivatives of ω we replace r by p. Furthermore, by (5.)

$$\frac{\partial \omega_{pk}}{\partial x_i} = \begin{bmatrix} ip \\ k \end{bmatrix} + \begin{bmatrix} ik \\ p \end{bmatrix}, \qquad \frac{\partial \omega_{pk}}{\partial x_h} = \begin{bmatrix} hp \\ k \end{bmatrix} + \begin{bmatrix} hk \\ p \end{bmatrix},$$

hence

$$\begin{bmatrix} pi \\ k \end{bmatrix} - \frac{\partial \omega_{pk}}{\partial x_i} = -\begin{bmatrix} ik \\ p \end{bmatrix}, \qquad \frac{\partial \omega_{pk}}{\partial x_h} - \begin{bmatrix} ph \\ k \end{bmatrix} = \begin{bmatrix} hk \\ p \end{bmatrix},$$

which gives us

$$= \frac{\partial \begin{bmatrix} gh \\ k \end{bmatrix}}{\partial x_i} - \frac{\partial \begin{bmatrix} gi \\ k \end{bmatrix}}{\partial x_h} + \sum_p \left(\begin{Bmatrix} gi \\ p \end{Bmatrix} \begin{bmatrix} hk \\ p \end{bmatrix} - \begin{Bmatrix} gh \\ p \end{Bmatrix} \begin{bmatrix} ik \\ p \end{bmatrix} \right).$$

We will denote this expression by $(gkhi)$[12], such that when we rewrite $\begin{Bmatrix} gi \\ p \end{Bmatrix}$, $\begin{Bmatrix} gh \\ p \end{Bmatrix}$ using (7.), we find

$$(13.)\qquad (gkhi) = \frac{\partial \begin{bmatrix} gh \\ k \end{bmatrix}}{\partial x_i} - \frac{\partial \begin{bmatrix} gi \\ k \end{bmatrix}}{\partial x_h} + \sum_{\alpha\beta} \frac{E_{\alpha\beta}}{E} \left(\begin{bmatrix} gi \\ \alpha \end{bmatrix} \begin{bmatrix} hk \\ \beta \end{bmatrix} - \begin{bmatrix} gh \\ \alpha \end{bmatrix} \begin{bmatrix} ik \\ \beta \end{bmatrix} \right),$$

which is by performing the derivatives equal to

$$(14.)\qquad \begin{cases} (gkhi) = \dfrac{1}{2} \left(\dfrac{\partial^2 \omega_{gi}}{\partial x_h \partial x_k} + \dfrac{\partial^2 \omega_{hk}}{\partial x_g \partial x_i} - \dfrac{\partial^2 \omega_{gh}}{\partial x_i \partial x_k} - \dfrac{\partial^2 \omega_{ik}}{\partial x_g \partial x_h} \right) \\ \qquad + \displaystyle\sum_{\alpha\beta} \frac{E_{\alpha\beta}}{E} \left(\begin{bmatrix} gi \\ \alpha \end{bmatrix} \begin{bmatrix} hk \\ \beta \end{bmatrix} - \begin{bmatrix} gh \\ \alpha \end{bmatrix} \begin{bmatrix} ik \\ \beta \end{bmatrix} \right). \end{cases}$$

With this definition, the integrability conditions (12.) take the following form:

$$(15.)\qquad (\alpha\delta\beta\gamma)' = \sum_{ghik} (ghik)\, u_\alpha^g u_\beta^h u_\gamma^i u_\delta^k.$$

For this system of equations we have the same result as we found for (12.), they contain all necessary integrability conditions of (9.), but not (9.) itself, as they imply the more general system (9′.) subject to (11.). Hence they cannot fully replace the integrability conditions of (1.), but are a necessary consequence.

The coefficients of $(ghik)$ have the following properties. Exchanging i and h in (13.) we find

$$(16^a.)\qquad (gkih) = -(gkhi).$$

Exchanging g and k in (14.) we find

$$(16^b.)\qquad (kghi) = -(gkhi).$$

Exchanging g and i, and k and h, the right hand side of (14.) does not change, as $E_{\beta\alpha} = E_{\alpha\beta}$, and we find

$$(ihkg) = (gkhi)$$

or

$$(16^c.)\qquad (higk) = (gkhi).$$

Finally we find

$$(16^d.)\qquad (gkhi) + (ghik) + (gikh) = 0.$$

With these four formulas, the number of truly different equations contained in (15.) is easily determined.

Because of (a.) and (b.) we can discard all cases where either $\alpha = \delta$ or $\beta = \gamma$, as well as those which follow for exchanging α with δ or β with γ. Hence the pair $\alpha\delta$, and similarly $\beta\gamma$, may be enumerated by two distinct, ordered elements of $1, 2, \ldots, n$, which gives us a total of $\frac{n(n-1)}{2} = n_2$ possible combinations.

[12] $(gkhi)$ really is the Riemann curvature tensor with all its indices lowered.

The expressions $(\alpha\delta\beta\gamma)'$ consist of three groups: the group where $\alpha\delta = \beta\gamma$ consisting of n_2 elements, a second of $\frac{n_2(n_2-1)}{2}$ expressions where $\alpha\delta$ does not equal $\beta\gamma$, and finally a third with an equal amount of expressions as the second, where $\alpha\delta$ and $\beta\gamma$ are exchanged. The third group may be discarded because of (c.). Hence we retain, because of (a.), (b.), and (c.) just $\frac{n_2(n_2+1)}{2}$ expressions $(\alpha\delta\beta\gamma)'$.

As every expression $(\alpha\delta\beta\gamma)'$ in which all of α, β, γ, and δ are distinct gives us a form of (d.), we see that

$$n_4 = \frac{1}{24} n(n-1)(n-2)(n-3)$$

expressions may be rewritten in terms of others, and just

$$\frac{n_2(n_2+1)}{2} - n_4 = \frac{n^2(n^2-1)}{12}$$

distinct expressions $(\alpha\delta\beta\gamma)'$ remain; this is the number of equations contained in (15.) that do not directly follow from each other.

The number of equations contained in (10.) equals $\frac{n(n+1)}{2}$, so adding these to the number of equations in (15.) we find

$$\frac{n(n+1)}{2} + \frac{n^2(n^2-1)}{12} = n^2 + n + \frac{(n+2)(n+1)n(n-3)}{12}$$

equations and therefore for $n = 2$ the number of equations in (10.) and (15.) is less than the number $n^2 + n$ of unknowns, being the variables x_i and their first derivatives u^i_α, for $n = 3$ they are equal, and for $n > 3$ the number of equations is greater than the number of unknowns.

6.

Just as the equations (10.) are transformation relations pertaining to (1.) or (2.), so (15.) contains the transformation relations of a form of order four. Before we continue to the actual construction of this form we will first show that the transformation relations of a more general form together with (9.) give us the transformation relations of a new form, which has an order equal to the order of the original form plus one.

Let

$$G_\mu = \sum_{i_1,\ldots,i_\mu} (i_1 i_2 \ldots i_\mu) \partial_1 x_{i_1} \partial_2 x_{i_2} \ldots \partial_\mu x_{i_\mu}$$

be a μ-linear form in the differentials $\partial_1 x$, $\partial_2 x$, $\ldots$, $\partial_\mu x$, derived from the coefficients of F. Let G'_μ denote its transform, with

$$(\alpha_1 \alpha_2 \ldots \alpha_\mu)' = \sum_{i_1 \ldots i_\mu} (i_1 i_2 \ldots i_\mu)\, u^{i_1}_{\alpha_1} u^{i_2}_{\alpha_2} \ldots u^{i_\mu}_{\alpha_\mu},$$

where

$$u^i_\alpha = \frac{\partial x_i}{\partial x_\alpha}.$$

Differentiating this expression to x'_α, we find

$$\frac{\partial(\alpha_1\alpha_2\ldots\alpha_\mu)'}{\partial x'_\alpha} = \sum_{i_1\ldots i_\mu} \frac{\partial(i_1 i_2\ldots i_\mu)}{\partial x_i}\, u^i_\alpha u^{i_1}_{\alpha_1}\ldots u^{i_\mu}_{\alpha_\mu} + P,$$

$$P = \sum_{\lambda i_2\ldots i_\mu} (\lambda i_2\ldots i_\mu)\, \frac{\partial^2 x_\lambda}{\partial x'_\alpha \partial x'_{\alpha_1}} u^{i_2}_{\alpha_2}\ldots u^{i_\mu}_{\alpha_\mu}$$

$$+ \sum_{i_1\lambda\ldots i_\mu} (i_1\lambda\ldots i_\mu)\, u^{i_1}_{\alpha_1} \frac{\partial^2 x_\lambda}{\partial x'_\alpha \partial x'_{\alpha_2}} \ldots u^{i_\mu}_{\alpha_\mu} + \ldots.$$

If we now replace all second derivatives in P by

$$\frac{\partial^2 x_\lambda}{\partial x'_\alpha \partial x'_{\alpha_s}} = \sum_r \begin{Bmatrix}\alpha\alpha_s\\ r\end{Bmatrix}' u^\lambda_r - \sum_{i i_s} \begin{Bmatrix} i i_s\\ \lambda\end{Bmatrix} u^i_\alpha u^{i_s}_{\alpha_s}$$

then we can write P as a difference $U - V$ where

$$U = \sum_r \begin{Bmatrix}\alpha\alpha_1\\ r\end{Bmatrix}' \sum_{\lambda i_2\ldots i_\mu} (\lambda i_2\ldots i_\mu)\, u^\lambda_r u^{i_2}_{\alpha_2}\ldots u^{i_\mu}_{\alpha_\mu}$$

$$+ \sum_r \begin{Bmatrix}\alpha\alpha_2\\ r\end{Bmatrix}' \sum_{i_1\lambda\ldots i_\mu} (i_1\lambda\ldots i_\mu)\, u^{i_1}_{\alpha_1} u^\lambda_r \ldots u^{i_\mu}_{\alpha_\mu} + \ldots$$

$$= \sum_r \left[\begin{Bmatrix}\alpha\alpha_1\\ r\end{Bmatrix}' (r\alpha_2\ldots\alpha_\mu)' + \begin{Bmatrix}\alpha\alpha_2\\ r\end{Bmatrix}' (\alpha_1 r\ldots\alpha_m u)' + \ldots \right]$$

and

$$V = \sum_{i\ldots i_\mu} u^i_\alpha u^{i_1}_{\alpha_1}\ldots u^{i_\mu}_{\alpha_\mu} \sum_\lambda \left[\begin{Bmatrix} i i_1\\ \lambda\end{Bmatrix} (\lambda i_2\ldots i_\mu) + \begin{Bmatrix} i i_2\\ \lambda\end{Bmatrix} (i_1\lambda\ldots i_\mu) + \ldots \right].$$

Substituting these expressions in P and bringing U to the left hand side, we arrive at the following statement:

Under the condition that all integrability conditions are satisfied, every system of transformation relations of order μ

$$(17^a.)\qquad (\alpha_1\alpha_2\ldots\alpha_\mu)' = \sum_{i_1\ldots i_\mu} (i_1 i_2\ldots i_\mu)\, u^{i_1}_{\alpha_1} u^{i_2}_{\alpha_2}\ldots u^{i_\mu}_{\alpha_\mu}$$

yields a new system of transformation relations of order $\mu + 1$

$$(17^b.)\qquad (\alpha\alpha_1\ldots\alpha_\mu)' = \sum_{i\ldots i_\mu} (i i_1\ldots i_\mu) u^i_\alpha u^{i_1}_{\alpha_1}\ldots u^{i_\mu}_{\alpha_\mu},$$

where we define

$$(17^c.)\qquad (i i_1\ldots i_\mu) = \frac{\partial(i_1 i_2\ldots i_\mu)}{\partial x_i} - \sum_\lambda \left[\begin{Bmatrix} i i_1\\ \lambda\end{Bmatrix} (\lambda i_2\ldots i_\mu) + \begin{Bmatrix} i i_2\\ \lambda\end{Bmatrix} (i_1\lambda\ldots i_\mu) + \ldots \right]$$

and use a similar definition for the transformed form. [13]

Using this statement we can from an equation $G_\mu = G'_\mu$ construct a sequence of similar equations $G_{\mu+1} = G'_{\mu+1}$, $G_{\mu+2} = G'_{\mu+2}$, ..., until we arrive at identical relations or relations composed of earlier ones.

This in fact occurs for the form F itself. If we would take $(i_1 i_2) = \omega_{i_1 i_2}$, then

$$\begin{aligned}(ii_1i_2) &= \frac{\partial \omega_{i_1 i_2}}{\partial x_i} - \sum_\lambda \left[\begin{Bmatrix} ii_1 \\ \lambda \end{Bmatrix} \omega_{\lambda i_2} + \begin{Bmatrix} ii_2 \\ \lambda \end{Bmatrix} \omega_{i_1 \lambda} \right] \\ &= \frac{\partial \omega_{i_1 i_2}}{\partial x_i} - \begin{bmatrix} ii_1 \\ i_2 \end{bmatrix} - \begin{bmatrix} ii_2 \\ i_1 \end{bmatrix}\end{aligned}$$

which equals zero by equation (5.). [14]

7.

We will now construct the form G_4 which has (15.) as transformation relation:

$$(18.) \qquad (\alpha\alpha_1\alpha_2\alpha_3)' = \sum_{i\ldots i_3} (ii_1i_2i_3) \frac{\partial x_i}{\partial x'_\alpha} \frac{\partial x_{i_1}}{\partial x'_{\alpha_1}} \frac{\partial x_{i_2}}{\partial x'_{\alpha_2}} \frac{\partial x_{i_3}}{\partial x'_{\alpha_3}},$$

and by (16.) we have

$$(19.) \qquad \begin{cases} (ii_1i_3i_2) = -(ii_1i_2i_3), \quad (i_1ii_2i_3) = -(ii_1i_2i_3), \quad (i_2i_3ii_1) = (ii_1i_2i_3), \\ \qquad\qquad (ii_1i_2i_3) + (ii_2i_3i_1) + (ii_3i_1i_2) = 0. \end{cases}$$

We multiply equation (18.) with $\partial x'_\alpha \, \delta x'_{\alpha_1} \, Dx'_{\alpha_2} \, \Delta x'_{\alpha_3}$, where we consider ∂, δ, D, and Δ to be independent differentials, and define

$$\sum_{i\ldots i_3} (ii_1i_2i_3) \, \partial x_i \, \delta x_{i_1} \, Dx_{i_2} \, \Delta x_{i_3} = G_4;$$

then from (18.) we obtain by summing over over all values of α that

$$G'_4 = G_4.$$

Here G_4 is a four-linear form in the variables ∂x, δx, Dx, and Δx, all subject to the same linear substitutions[15]. Because of the properties of the coefficients, this form has a number of interesting properties.

If we exchange i and i_1, which does not change G_4 as we sum over all values, and replace $(i_1ii_2i_3)$ by $-(ii_1i_2i_3)$, then it follows that G_4 only changes sign

[13] This really is the covariant derivative ∇T of a tensor T: on a local coordinate patch of a manifold M we have, if $T(x) = \sum_{a_1,\ldots,a_k} T_{a_1\ldots a_k}(x) \, \mathrm{d}x^{a_1} \otimes \ldots \otimes \mathrm{d}x^{a_k}$ is a covariant tensor of order k, then the covariant derivative ∇T of T is a covariant tensor of order $k+1$ and is given in local coordinates by $\nabla T(x) = \sum_{a,a_1,\ldots,a_k} \left(\frac{\partial T(x)}{\partial x_a} + \sum_b \left[\Gamma^b_{aa_1}(x) T_{ba_2\ldots a_k}(x) + \Gamma^b_{aa_2}(x) T_{a_1 b\ldots a_k}(x) + \ldots + \Gamma^b_{aa_k}(x) T_{a_1 a_2 \ldots b}(x) \right] \right) \mathrm{d}x^a \otimes \mathrm{d}x^{a_1} \otimes \ldots \otimes \mathrm{d}x^{a_k}$ (compare with (17^c.)). Hence (17^b.) is an expression of the fact that the covariant derivative of a covariant tensor is again a covariant tensor.

[14] Which shows that the metric is covariantly constant with respect to this (covariant) derivative.

[15] Under changing of variables.

when ∂ and δ are exchanged. Now if we only sum over the distinct ii_1 terms[16], then

$$G_4 = \sum (ii_1i_2i_3)\,(\partial x_i\,\delta x_{i_1} - \partial x_{i_1}\,\delta x_i) Dx_{i_2}\,\Delta x_{i_3}.$$

The same thing happens for D and Δ when i_2 and i_3 are exchanged, and it follows that

$$(20.)\qquad G_4 = {}'\sum (ii_1i_2i_3)\,(\partial x_i\,\delta x_{i_1} - \partial x_{i_1}\,\delta x_i)\,(Dx_{i_2}\,\Delta x_{i_3} - Dx_{i_3}\,\Delta x_{i_2}),$$

where ${}'\sum$ denotes summing over the $\frac{n(n-1)}{2}$ distinct values of ii_1, and similarly for i_2i_3.

If we exchange i and i_2, as well as i_1 and i_3, and use the equation $(i_2i_3ii_1) = (ii_1i_2i_3)$, then G_4 remains the same, while $\partial x_i\,\delta x_k - \partial x_k\,\delta x_i$ and $Dx_i\,\Delta x_k - Dx_k\,\Delta x_i$ are exchanged. Even so, we may not set $D = \partial$ and $\Delta = \delta$, because then the coefficients $(ii_1i_2i_3)$ and $(i_2i_1ii_3)$, or equivalently $(ii_1i_2i_3)$ and $(ii_3i_2i_1)$, are summed together which prohibits deriving (18.) from $G_4 = G_4'$.

Finally, cyclically permuting δ, D, and Δ yields the forms

$$H_4 = \sum (ii_2i_3i_1)\,\partial x_i\,\delta x_{i_1}\,Dx_{i_2}\,\Delta x_{i_3} = \sum (ii_1i_2i_3)\,\partial x_i\,Dx_{i_1}\,\Delta x_{i_2}\,\delta x_{i_3},$$
$$J_4 = \sum (ii_3i_1i_2)\,\partial x_i\,\delta x_{i_1}\,Dx_{i_2}\,\Delta x_{i_3} = \sum (ii_1i_2i_3)\,\partial x_i\,\Delta x_{i_1}\,\delta x_{i_2}\,Dx_{i_3},$$

for which, using (19.), we have

$$G_4 + H_4 + J_4 = 0.$$

8.

Using the coefficients $(i_1i_2i_3i_4)$ and applying the technique of section 6. to G_4 we obtain a new set of coefficients $(ii_1i_2i_3i_4)$ from which we build a five-linear form G_5 using the same technique as in section 7. We can again do the same with G_5 to construct a six-linear form G_6, from G_6 a seven-linear form G_7, etc., until we encounter forms whose coefficients vanish or reduce to those of previous forms. This gives us the system of equations

$$(A.)\qquad F = F',\quad G_4 = G_4',\quad G_5 = G_5',\quad \ldots,$$

and from the previous considerations we see that the validity equations on the right are a necessary consequence of the validity of those on the left. [17]

For the equation $F = F'$ to hold it is both necessary and sufficient that the transformation relations

$$(F.)\qquad (\alpha_1\alpha_2)' = \sum_{i_1i_2} (i_1i_2)\,u^{i_1}_{\alpha_1}u^{i_2}_{\alpha_2}$$

for $(ik) = \omega_{ik}$, as well as the integrability conditions

$$(B.)\qquad \frac{\partial u^i_\alpha}{\partial x'_\beta} = \frac{\partial u^i_\beta}{\partial x'_\alpha}$$

[16] That is, only sum over all i and i_1 which satisfy $i < i_1$.

[17] From $F = F'$ we found that necessarily $G_4 = G_4'$ by (15.), just as $G_5 = G_5'$ is a direct consequence of $G_4 = G_4'$ by (17^b.), ...

are satisfied. Because if both conditions are satisfied, then there exist n functions $v_1, v_2, \ldots, v_n$ such that the equations $(F.)$ hold for

$$u_\alpha^i = \frac{\partial v_i}{\partial x'_\alpha}$$

and the substitution $x_1 = v_1,\ x_2 = v_2,\ \ldots,\ x_n = v_n$ then gives the desired transformation of F into F'. [18]

For the equation $G_4 = G'_4$, we need the transformation relations

$$(G_4.) \qquad (\alpha_1\alpha_2\alpha_3\alpha_4)' = \sum_{i_1\ldots i_4} (i_1 i_2 i_3 i_4)\, u_{\alpha_1}^{i_1} u_{\alpha_2}^{i_2} u_{\alpha_3}^{i_3} u_{\alpha_4}^{i_4}$$

and the integrability conditions $(B.)$ to be satisfied.

In the same fashion the integrability conditions and

$$(G_5.) \qquad (\alpha\alpha_1 \ldots \alpha_4)' = \sum_{ii_1\ldots i_4} (ii_1 \ldots i_4)\, u_\alpha^i u_{\alpha_1}^{i_1} \ldots u_{\alpha_4}^{i_4}$$

must necessarily be satisfied for $G_5 = G'_5$ to hold, etc..

With this observation, the transformation of F into F' is impossible when the transformation relations $(F.)$, $(G_4.)$, $(G_5.)$, $(G_6.)$, etc. cannot simultaneously be satisfied, regardless of the necessary integrability conditions $(B.)$.

The question now remains whether or not the integratbility conditions $(B.)$ are superfluous when the transformation relations are all satisfied.

9.

To answer this question and avoid unnecessarily long and tedious calculations, we will restrict ourselves in this section to a particular case, which nevertheless allows for direct generalisation, and then find an appropriate answer to our question above.

Suppose that 1) the unknowns x and u are completely and uniquely determined by for example the system of equations $(G_4.)$, and that 2) their values satisfy the next system of equations $(G_5.)$.

Taking the derivative of $(G_4.)$ with respect to x'_α, we obtain

$$(G'_4.) \qquad \begin{cases} \dfrac{\partial(\alpha_1\alpha_2\alpha_3\alpha_4)'}{\partial x'_\alpha} = \displaystyle\sum_{i\ldots i_4} \frac{\partial(i_1 i_2 i_3 i_4)}{\partial x_i} \frac{\partial x_i}{\partial x'_\alpha} u_{\alpha_1}^{i_1} \ldots u_{\alpha_4}^{i_4} + \Pi, \\ \Pi = \displaystyle\sum_{\lambda i_2 i_3 i_4} (\lambda i_2 i_3 i_4) \frac{\partial u_{\alpha_1}^\lambda}{\partial x'_\alpha} u_{\alpha_2}^{i_2} u_{\alpha_3}^{i_3} u_{\alpha_4}^{i_4} \\ \qquad + \displaystyle\sum_{i_1 \lambda i_3 i_4} (i_1 \lambda i_3 i_4)\, u_{\alpha_1}^{i_1} \frac{\partial u_{\alpha_2}^\lambda}{\partial x'_\alpha} u_{\alpha_3}^{i_3} u_{\alpha_4}^{i_4} + \ldots. \end{cases}$$

From our first assumption it follows that the values of all derivatives

$$\frac{\partial x_i}{\partial x'_\alpha}, \quad \frac{\partial u_{\alpha_s}^\lambda}{\partial x'_\alpha}$$

[18] See [DK2004II], Lemma 8.2.6 (Poincaré) for the existence of such $v_1, \ldots, v_n$.

are uniquely determined, and therefore that $(G_4'.)$ contains a number of equations equal to the number of these unknowns, which are all independent and do not contradict each other.

If we now let

$$\frac{\partial x_i}{\partial x'_\alpha} - u^i_\alpha = \binom{i}{\alpha}$$

$$\frac{\partial u^\lambda_{\alpha_s}}{\partial x'_\alpha} + \sum_{ii_s} \begin{Bmatrix} ii_s \\ \lambda \end{Bmatrix} u^i_\alpha u^{i_s}_{\alpha_s} - \sum_r \begin{Bmatrix} \alpha\alpha_s \\ r \end{Bmatrix}' u^\lambda_r = \binom{\lambda}{\alpha\alpha_s},$$

then $\binom{i}{\alpha} = 0$, $\binom{\lambda}{\alpha\alpha_s} = 0$ are the equations that are necessary to derive the equations $(G_5.)$ from $(G_4'.)$ in the same way as we did in section 6..

We can get rid of all derivatives in $(G_4'.)$ using these equations and retain on the right hand side a part U_4 that is linear in $\binom{i}{\alpha}$ and $\binom{\lambda}{\alpha\alpha_s}$, which is different from the original right hand side only through replacing $\frac{\partial x_i}{\partial x'_\alpha}$ by $\binom{i}{\alpha}$, and $\frac{\partial u^\lambda_{\alpha_s}}{\partial x'_\alpha}$ by $\binom{\lambda}{\alpha\alpha_s}$. The other terms cancel against the left hand side, where we use that the equations in $(G_5.)$ are satisfied by x and u. Hence the equations $(G_4'.)$ reduce to

$$U_4 = 0,$$

which is the same as $(G_4'.)$ with all the unknowns replaced by $\binom{i}{\alpha}$, $\binom{\lambda}{\alpha\alpha_s}$ and all other terms removed.

As noted before, this system of equations contains a number of equations equal to the number of unknowns $\binom{i}{\alpha}$, $\binom{\lambda}{\alpha\alpha_s}$ and the equations are all independent from each other. Hence they can only be satisfied by $\binom{i}{\alpha} = 0$ and $\binom{\lambda}{\alpha\alpha_s} = 0$, that is, if for all i and α

$$u^i_\alpha = \frac{\partial x_i}{\partial x'_\alpha}.$$

If therefore the equations $(G_4.)$ uniquely determines all unknows, and the values of these unknowns satisfy the equations $(G_5.)$ following from $(G_4.)$, then the integrability conditions are satisfied and hence the equations $(B.)$ become superfluous.

10.

After section 9. it is clear without further calculations that the same conclusion is valid when all unknowns have completely and uniquely been determined by the equations $(G_p.)$, $(G_q.)$, $(G_r.)$, ... and furthermore satisfy $(G_{p+1}.)$, $(G_{q+1}.)$, $(G_{r+1}.)$, ... which are derived using the techniques of section 6.. Then from every system of equations $(G_s.)$ together with $(G_{s+1}.)$ we obtain a new set of equations $U_s = 0$, which gives us a sequence of equations

$$U_p = 0, \quad U_q = 0, \quad U_r = 0, \quad \dots$$

from which we must make a selection consisting of an equal number of equations as unknowns, which form a system with nonvanishing determinant.[19]

First of the equations which we can use to determine x and u is $(F.)$ which distinguishes itself from the other $(G_4.)$, $(G_5.)$, ... equations by the fact that the equations given by their derivative are always satisfied, as we saw in section 6.: for $(i_1 i_2) = \omega_{i_1 i_2}$, $(i i_1 i_2) = 0$. Hence we arrive at the following statement:

From the original differential expression F we first, using sections 5. and 7., derive the form G_4 and then using section 6. the sequence of forms G_5, G_6, etc.. When the transformation relations determined by the equations

$$F = F', \quad G_4 = G_4', \quad G_5 = G_5', \quad \dots, \quad G_p = G_p'$$

uniquely determine x and u, and these values satisfy the transformation relations given by

$$G_{p+1} = G_{p+1}',$$

then the necessary integrability conditions for the transformation of F into F' are satisfied, and we have

$$u_\alpha^i = \frac{\partial x_i}{\partial x_\alpha'}.$$

This statement, which in all cases where it may be applied makes the integrability conditions superfluous, opens up a new area of applicability for the theory of algebraic invariants.

It enables us to consider F, G_4, G_5, ... just as homogeneous algebraic forms in the variables ∂x, $\partial_1 x$, ..., where it is no longer necessary to consider these as differentials, provided these variables are all subject to the same linear substitution.[20]

Under these conditions we have, which follows directly from algebra, that we can continue the sequence of forms F, G_4, G_5, ... until we have constructed n absolute invariants and an equal amount of corresponding forms Ψ_s such that the invariants are independent as functions of x and the Ψ_s do not depend on each other. Then if we continue the sequence F, G_4, G_5, ... one term further, we find a complete system of invariants I, I_1, ... which are in relation to the coefficients of the forms in the sequence, mutually independent.

With these invariants the possibility of transforming F into F' exists if and only if the equations

$$I' = r^\lambda I, \quad I_1' = r^{\lambda_1} I_1, \quad \dots,$$

where r is the determinant of the substitution and the exponents λ are constant.

It is therefore appropriate to *call these invariants I, I_1, ... belonging to the sufficiently long, but not longer than necessary sequence F, G_4, G_5, ... the complete system of invariants of the differential expression F.*

[19] And hence uniquely determine the unknowns such that $(B.)$ is satisfied.
[20] Described by the matrix $\left(u_\alpha^i\right)_{i,\alpha}$ with determinant r.

Let U_1, U_2, …, U_n denote the variables in the original form, as well as V_1, V_2, …, V_n the variables of the transformed form, then the original substitution

$$\partial x_i = \sum_\alpha u^i_\alpha \partial x'_\alpha$$

yields

$$V_\alpha = \sum_i u^i_\alpha U_i$$

as transpose. Then the equations between the corresponding forms become

$$\Psi'_s(V_1, V_2, \ldots) = r^{\mu_s}\,\Psi_s(U_1, U_2, \ldots),$$

for $s = 1, 2, \ldots, n$ and with μ constant. It follows that if for a given function Ω depending on x we set

$$U_i = \frac{\partial \Omega}{\partial x_i}$$

the equations are

$$\Psi'_s\Big(\frac{\partial \Omega}{\partial x'_1}, \frac{\partial \Omega}{\partial x'_2}, \ldots\Big) = r^{\mu_s}\,\Psi_s\Big(\frac{\partial \Omega}{\partial x_1}, \frac{\partial \Omega}{\partial x_2}, \ldots\Big).$$

We call an expression

$$\Psi_s\Big(\frac{\partial \Omega}{\partial x_1}, \frac{\partial \Omega}{\partial x_2}, \ldots\Big)$$

containing an arbitrary function Ω, an associated form of the differential expression F.

In a similar fashion the covariants of system F, G_4, G_5, … give n equations

$$\Phi'_s(\partial x'_1, \partial x'_2, \ldots) = r^{\nu_s}\,\Phi_s(\partial x_1, \partial x_2, \ldots),$$

and *we call the differential expressions*

$$\Phi_s(\partial x_1, \partial x_2, \ldots)$$

the covariants of the differential expression F.

11.

To make the contents of the statement from the previous section more precise, it must be emphasized that the conditions of this statement cannot always be fulfilled. For example, consider the question of whether or not two given surfaces can be mapped onto each other without distorting them, [21] then we must investigate the equation $F = F'$ for $n = 2$. Now even if $F = F'$ can be satisfied for a certain transformation, the transformation relations given by $F = F'$, $G_4 = G'_4$, $G_5 = G'_5$, … may still fail to uniquely determine x and u, no matter how many equations are considered. This is the case when, for example,

[21] That is, via an isometry with respect to the metrics on the surfaces.

the planes can be translated within themselves, changing the original variables while keeping the new variables fixed. [22]

Now it is easy to determine the conditions for this case to occur.

The domain of the variables $x_1, x_2, \ldots, x_n$ is called translatable[23], whenever it is possible to find a substitution of variables, such that the transformation of F does not uniquely determine this substitution. It is clear that the conditions for this to occur consist of identities between invariants and forms of F, which under the conditions of the previous section should be independent functions of x and u.

In this case the integrability conditions are not necessarily fulfilled, because even if all transformation relations are satisfied, the terms $\begin{pmatrix} i \\ \alpha \end{pmatrix}$, $\begin{pmatrix} \lambda \\ \alpha\alpha_s \end{pmatrix}$ need not be zero, because they are no longer uniquely determined.

For all other cases we have the pleasant and a priori unexpected result that for the possibility of transforming F to F' necessary and sufficient conditions can be phrased as equations between invariants [24] and that the forms and covariants belonging to the transformation problem can be considered and treated in a purely algebraic way.

12.

As an example of the discussed theory, we now treat the case $n = 3$, which shows some remarkable simplifications.

In equation (20.) we take for ii_1 and i_2i_3 only the unique pairs in $1\,2\,3$, so for example just $2\,3$, $3\,1$, and $1\,2$. Then the transformation relations (18.) of G_4 can be written as

$$(\beta_1\beta_2\alpha_1\alpha_2)' = {}'\sum(i_1i_2k_1k_2)\,(u^{i_1}_{\beta_1}u^{i_2}_{\beta_2} - u^{i_1}_{\beta_2}u^{i_2}_{\beta_1})(u^{k_1}_{\gamma_1}u^{k_2}_{\gamma_2} - u^{k_1}_{\gamma_2}u^{k_2}_{\gamma_1}).$$

We now define four numbers β, γ, i, and k by requiring that $\beta\beta_1\beta_2$, $\gamma\gamma_1\gamma_2$, ii_1i_2, and kk_1k_2 form positive permutations of the numbers $1\,2\,3$, such that

$$\begin{aligned} &\beta = 1, \quad \beta_1 = 2, \quad \beta_2 = 3, \\ \text{or}\quad &\beta = 2, \quad \beta_1 = 3, \quad \beta_2 = 1, \\ \text{or}\quad &\beta = 3, \quad \beta_1 = 1, \quad \beta_2 = 2. \end{aligned}$$

Conversely we see that from β, we can uniquely determine β_1 and β_2.

Let r^i_α be the minor of the determinant r [25] and since the values of i_1i_2 and k_1k_2 are completely determined by the above defined i and k, we put

$$(i_1i_2k_1k_2) = A_{ik},$$

[22] Consider for example the two planes $\{(x, y, 0) \in \mathbb{R}^3 | x, y \in \mathbb{R}\}$ and $\{(x, y, 1) \in \mathbb{R}^3 | x, y \in \mathbb{R}\}$ with metric induced by the Euclidean metric in $\mathbb{R}^3$. Then the transformation taking the first plane into the second is not at all uniquely determined: for any fixed $(x_0, y_0) \in \mathbb{R}^2$ the map $(x, y, 0) \mapsto (x + x_0, y + y_0, 1)$ is an isometry.

[23] *"verschiebbares"*

[24] *"[. . .], wenn dieser Ausdruck zur Bezeichnung der gleichen Formverhältnisse wie in der Algebra angewandt wird"*

[25] Recall that r is the determinant of the matrix $\left(\frac{\partial x_i}{\partial x'_\alpha}\right)_{i,\alpha} = (u^i_\alpha)_{i,\alpha}$.

such that by (14.) we have

$$(a.)\qquad \begin{cases} A_{11} = \dfrac{1}{2}\left(2\dfrac{\partial^2\omega_{23}}{\partial x_2 \partial x_3} - \dfrac{\partial^2\omega_{22}}{\partial x_3^2} - \dfrac{\partial^2\omega_{33}}{\partial x_2^2}\right) \\ \qquad + \displaystyle\sum_{\alpha\beta} \frac{E_{\alpha\beta}}{E}\left(\begin{bmatrix}23\\ \alpha\end{bmatrix}\begin{bmatrix}23\\ \beta\end{bmatrix} - \begin{bmatrix}22\\ \alpha\end{bmatrix}\begin{bmatrix}33\\ \beta\end{bmatrix}\right) \\ A_{23} = \dfrac{1}{2}\left(\dfrac{\partial^2\omega_{23}}{\partial x_1^2} + \dfrac{\partial^2\omega_{11}}{\partial x_2 \partial x_3} - \dfrac{\partial^2\omega_{12}}{\partial x_1 \partial x_3} - \dfrac{\partial^2\omega_{13}}{\partial x_1 \partial x_2}\right) \\ \qquad + \displaystyle\sum_{\alpha\beta} \frac{E_{\alpha\beta}}{E}\left(\begin{bmatrix}11\\ \alpha\end{bmatrix}\begin{bmatrix}23\\ \beta\end{bmatrix} - \begin{bmatrix}12\\ \alpha\end{bmatrix}\begin{bmatrix}13\\ \beta\end{bmatrix}\right), \end{cases}$$

from which by cyclic permutations of 1 2 3 the other components follow.

Then the transformation relations of G_4 become: [26]

$$(b.)\qquad A'_{\beta\gamma} = \sum_{ik} A_{ik} r^i_\beta r^k_\gamma.$$

In a similar way the transformation relations of $F = F'$ can be replaced by:

$$(c.)\qquad E'_{\beta\gamma} = \sum_{ik} E_{ik} r^i_\beta r^k_\gamma.$$

Furthermore, because of (19.) we have

$$(d.)\qquad A_{ki} = A_{ik}$$

and A_{ik} changes sign if we exchange i_1 with i_2 or k_1 with k_2.

The equations $(b.)$ and $(c.)$ are nothing but transformation relations for the simultaneous transformations of the to F associated quadratic forms

$$\Gamma = \sum_{ik} A_{ik} X_i X_k,$$
$$\Phi = \sum_{ik} E_{ik} X_i X_k$$

into

$$\Gamma' = \sum_{ik} A_{ik} \Xi_i \Xi_k,$$
$$\Phi' = \sum_{ik} E_{ik} \Xi_i \Xi_k$$

via the substitution

$$(e.)\qquad X_i = \sum_\beta r^i_\beta \Xi_\beta,$$

with determinant

$$R = r^2$$

[26]This is a consequence of the first equation of section 12. and definition of the r^i_α.

and inverse [27]

$$(e'.) \qquad \Xi_\beta = \sum_i \frac{1}{r} u^i_\beta X_i.$$

It is known that this transformation problem can be solved if four simultaneous invariants and three corresponding forms exist that are independent with respect to the variables X_i, Ξ_β and the coefficients of Γ and Φ. These yield three absolute invariants; if these are independent with respect to the variables x_1, x_2, x_3, then by the statement from section 10. it is only necessary to determine G_5 to solve the problem.

By section 6. we find

$$(gi_1i_2k_1k_2) = \frac{\partial(i_1i_2k_1k_2)}{\partial x_g} - \sum_\lambda \left[\begin{Bmatrix} gi_1 \\ \lambda \end{Bmatrix} (\lambda i_2k_1k_2) + \begin{Bmatrix} gi_2 \\ \lambda \end{Bmatrix} (i_1\lambda k_1k_2) + \begin{Bmatrix} gk_1 \\ \lambda \end{Bmatrix} (i_1i_2\lambda k_2) + \begin{Bmatrix} gk_2 \\ \lambda \end{Bmatrix} (i_1i_2k_1\lambda) \right].$$

In this summation we for the first term only need to take the values i and i_1 for λ, as $(i_2i_2k_1k_2) = 0$. Using similar considerations for the other three terms, the cases $\lambda = i$ and $= k$ are identical, we obtain

$$(gi_1i_2k_1k_2) = \frac{\partial(i_1i_2k_1k_2)}{\partial x_g} - (i_1i_2k_1k_2) \left[\begin{Bmatrix} gi_1 \\ i_1 \end{Bmatrix} + \begin{Bmatrix} gi_2 \\ i_2 \end{Bmatrix} + \begin{Bmatrix} gk_1 \\ k_1 \end{Bmatrix} + \begin{Bmatrix} gk_2 \\ k_2 \end{Bmatrix} \right]$$
$$+ \begin{Bmatrix} gi_1 \\ i \end{Bmatrix} (i_2ik_1k_2) + \begin{Bmatrix} gi_2 \\ i \end{Bmatrix} (ii_1k_1k_2) + \begin{Bmatrix} gk_1 \\ k \end{Bmatrix} (i_1i_2k_2k) + \begin{Bmatrix} gk_2 \\ k \end{Bmatrix} (i_1i_2kk_1).$$

Now as i_1i_2i, i_2ii_1, and k_1k_2k, k_2kk_1 together with ii_1i_2 and kk_1k_2 are positive permutations of 1 2 3, we obtain after adding $\begin{Bmatrix} gi \\ i \end{Bmatrix}$, $\begin{Bmatrix} gk \\ k \end{Bmatrix}$ in $\Big[\dots\Big]$:

$$(gi_1i_2k_1k_2) = \frac{\partial A_{ik}}{\partial x_g} - 2A_{ik} \sum_r \begin{Bmatrix} gr \\ r \end{Bmatrix}$$
$$+ \begin{Bmatrix} gi \\ i \end{Bmatrix} A_{ik} + \begin{Bmatrix} gi_1 \\ i \end{Bmatrix} A_{i_1k} + \begin{Bmatrix} gi_2 \\ i \end{Bmatrix} A_{i_2k} + \begin{Bmatrix} gk \\ k \end{Bmatrix} A_{ik} + \begin{Bmatrix} gk_1 \\ k \end{Bmatrix} A_{ik_1} + \begin{Bmatrix} gk_2 \\ k \end{Bmatrix} A_{ik_2}.$$

We denote this expression, which is completely determined by g, i, and k, by A_{gik}, such that

$$(gi_1i_2k_1k_2) = A_{gik},$$

and using the easily derived formula

$$2\sum_r \begin{Bmatrix} gr \\ r \end{Bmatrix} = \frac{1}{E}\frac{\partial E}{\partial x_g}$$

we find

$$(f.) \qquad A_{gik} = \frac{\partial A_{ik}}{\partial x_g} - \frac{A_{ik}}{E}\frac{\partial E}{\partial x_g} + \sum_\lambda \left[\begin{Bmatrix} g\lambda \\ i \end{Bmatrix} A_{\lambda k} + \begin{Bmatrix} g\lambda \\ k \end{Bmatrix} A_{i\lambda} \right].$$

[27] By Cramer's rule for the 3×3 matrix $\left(u^i_\alpha\right)_{i,\alpha}$ with determinant r.

This expression remains the same if we exchange i and k, and changes sign if we exchange i_1 with i_2 or k_1 with k_2. Using this notation, the equation $G_5 = G'_5$ becomes

$$A'_{\alpha\beta\gamma} = \sum_{gi_1i_2k_1k_2} A_{gik} u^g_\alpha u^{i_1}_{\beta_1} u^{i_2}_{\beta_2} u^{k_1}_{\gamma_1} u^{k_2}_{\gamma_2}$$

which simplifies to

$$(g.) \qquad A'_{\alpha\beta\gamma} = \sum_{gik} A_{gik} u^g_\alpha r^i_\beta r^k_\gamma$$

and leads to a peculiar result.

If we let for the forms Γ and Φ, U_1, U_2, U_3 be the original and V_1, V_2, V_3 be the new variables, such that the substitution corresponding to $(e.)$ becomes

$$(h.) \qquad V_\alpha = \sum_g r^g_\alpha U_g,$$

with inverse

$$(h'.) \qquad U_g = \sum_\alpha \frac{1}{r} u^g_\alpha V_\alpha,$$

then we obtain from $(g.)$ by multiplying with $V_\alpha \Xi_\beta \Xi_\gamma$ and summing over α, β, and γ:

$$\sum_{\alpha\beta\gamma} A'_{\alpha\beta\gamma} V_\alpha \Xi_\beta \Xi_\gamma = r \sum_{gik} A_{gik} U_g X_i X_k.$$

As $r = R^{\frac{1}{2}}$ is a power of the determinant of the substitution, and in both sums both the original variables as those of the associated forms occur [28], we arrive at the following statement:

To determine the possibility of the transformation of a quadratic differential expression

$$F = \sum \omega_{ik} \partial x_i \partial x_k$$

into

$$F' = \sum \omega'_{ik} \partial x'_i \partial x'_k,$$

use the coefficients of F to build three algebraic forms

$$\Gamma = \sum A_{ik} X_i X_k,$$
$$\Phi = \sum E_{ik} X_i X_k,$$
$$\Theta = \sum A_{gik} U_g X_i X_k,$$

with variables X and U, and similarly for F' the forms Γ', Φ', and Θ' with variables Ξ and V.

Now the conditions necessary and sufficient for the transformation of F into F' are precisely those for the existence of a substitution

$$X_i = \sum_\alpha r^i_\alpha \Xi_\alpha$$

[28] *"[...], so müssen dieselben entsprechende simultane Zwischenformen sein"*

which transforms Γ *into* Γ'*,* Φ *into* Φ'*, and via*

$$V_\alpha = \sum_i r_\alpha^i U_i$$

yields

$$\Theta' = R^{\frac{1}{2}}\Theta.$$

For this statement to be applicable, it is necessary that the six absolute associated forms and invariants of Γ *and* Φ *are mutually independent functions of the variables* x_1*,* x_2*,* x_3*,* U_1*,* U_2*,* U_3*.*

If this condition is satisfied, then the equations between the original and transformed covariants of Γ *and* Φ *yield the coefficients* $\frac{1}{r}u_\beta^i$ *of the inverse substitution, so as* r *is determined by the invariants, the equations yield* u_β^i *and in such a way that*

$$u_\beta^i = \frac{\partial x_i}{\partial x'_\beta}.$$

If the condition is not satisfied, then it follows that the domain of the variables x_1*,* x_2*, and* x_3 *can be translated into itself without changing* F*.*

From section 10. we know that under the conditions of this statement, the necessary and sufficient conditions for the transformation of F into F' are given by the simultaneous invariants of F, G_4, G_5 and their transforms. This result also shows why in the above statement, where we solve for the coefficients u_α^i or r_α^i of the substitution, we do not consider the integrability conditions between these coefficients. [29]

In the above statement it is demanded that the coefficients A_{gik} may be expressed in terms of the coefficients of Γ and Φ, and similarly for A'_{gik}, such that 1) Θ is a *Zwischenform* of Γ and Φ, 2) Θ' is a simultaneous *Zwischenform* of Γ' and Θ', and 3) $\Theta' = R^{\frac{1}{2}}\Theta$. From the exponent of R we see that these conditions may not be satisfied when A_{gik} is a rational function of A_{ik} and E_{ik}. On the other hand, these conditions are certainly satisfied whenever F' is not arbitrary but obtained from F through a direct substitution, such as the identity $x_i = x'_i$. Hence we find:

For every quadratic differential expression F*, we can find the coefficients* A_{gik} *defined in* $(f.)$ *as irrational functions of the coefficients of* Γ *and* Φ*, such that* Θ *is a simultaneous* Zwischenform *of* Γ *and* Φ*, which is related to its transform by* $\Theta' = R^{\frac{1}{2}}\Theta$*.*

Adding to this the equations $\Gamma' = \Gamma$, $\Phi' = \Phi$ and

$$\sum \pm r_1^1 r_2^2 r_3^3 = R,$$

then we retain 31 transformation relations from which we, even though the variables of Θ are subject to different substitutions, can solve the 9 substitution coefficients r_α^i in such a way that they have the invariant form

$$I' = R^\lambda I.$$

[29] *"Dieses Resultat setzt ebenso wie der obige Satz voraus, dass bei der Elimination der Substitutionscoefficienten* u_α^i *oder* r_α^i*, aus welcher die in Rede stehenden algebraischen Grundformen hervorgehen, auf Integrabilitätsbedingungen zwischen denselben keinerlei Rücksicht genommen werde."*

In the usual case where the transformations from Γ and Φ fully determine the substitution, the same result holds for Γ, Φ, and Θ. Then the number of independent invariants 1) from Γ and Φ alone is equal to 4, 2) from Γ, Φ, and Θ together equal to 22, so 3) the number of them which necessarily contain coefficients of Θ equals 18.

Using the terminology of section 10., for the current case the differential expression F has 22 invariants, 21 absolute invariants, 3 independent forms and an equal amount of covariants.

About the differential expression F given by the square of the line element in three dimensional space, which is not covered by the statement from this section, there is a treatment made by *Riemann* [30], for which Mr. *Dedekind* is treating the analytical backgrounds. [31]

3rd of Januari 1869.

[30] *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen.* Abh. der Göttinger Ges. d. W., 1867, Band XIII.

[31] *"[. . .], zu welcher Herr Dedekind die dort unterdrückten analytischen Entwicklungen in Aussicht gestellt hat."*

Bibliography

[Chr1869] E. B. Christoffel: *Ueber die Transformation der homogenen Differentialausdrücke zweiten Grades*, Reine Angewandte Mathematik 70 (1869), pages 46–70.

[Bou1947] N. Bourbaki: *Éléments de Mathématique VI*, Livre II: Algèbre, Chapitre II, Libraire Scientifique Hermann et C$^{\text{ie}}$ Paris (1947).

[Car1951] E. Cartan: *Géométrie des Espaces de Riemann*, deuxième edition, Gauthier-Villars, Paris (1951).

[Bou1953] N. Bourbaki: *Éléments de Mathématique XV*, Livre V: Espaces Vectoriels Topologiques, Chapitre I–II, Libraire Scientifique Hermann et C$^{\text{ie}}$ Paris (1953).

[Bou1955] N. Bourbaki: *Éléments de Mathématique XVIII*, Livre V: Espaces Vectoriels Topologiques, Chapitre III–V, Libraire Scientifique Hermann et C$^{\text{ie}}$ Paris (1955).

[Hus1965] T. Husain: *The Open Mapping and Closed Graph Theorems in Topological Vector Spaces*, Oxford Mathematical Monographs (1965).

[Ham1982] R. S. Hamilton: *The inverse function theorem of Nash and Moser*, Bulletin of the American Mathematical Society 7 (1982), pages 65–222.

[Wal1984] R. M. Wald: *General Relativity*, The University of Chicago Press (1984), ISBN 0226870332.

[DK2000] J. J. Duistermaat, J. A. C. Kolk: *Lie Groups*, Springer (2000), ISBN 3540152938.

[Mun2000] J. R. Munkres: *Topology* (second edition), Prentice Hall (2000), ISBN 0131784498.

[Dui2003] J. J. Duistermaat: *Functies en Reeksen*, lecture notes of Utrecht University (2003).

[DK2004I] J. J. Duistermaat, J. A. C. Kolk: *Multidimensional Real Analysis I*, Cambridge University Press (2004), ISBN 0521551145.

[DK2004II] J. J. Duistermaat, J. A. C. Kolk: *Multidimensional Real Analysis II*, Cambridge University Press (2004), ISBN 0521829259.

[Dui2006] J. J. Duistermaat: *Klassieke Mechanica*, lecture notes of Utrecht University (2006).

[Ban2008] E. P. van den Ban: *Riemannian geometry*, lecture notes of Utrecht University (2008).

[StAndrews] Biography of Elwin Bruno Christoffel, from `http://www-history.mcs.st-andrews.ac.uk/Biographies/Christoffel.html`, School of Mathematics and Statistics, University of St Andrews, Scotland, 1997.